

## PATENT ABSTRACTS OF JAPAN

(11)Publication number : 2000-155756

(43)Date of publication of application : 06.06.2000

(51)Int.Cl.

G06F 17/27  
G06F 17/30

(21)Application number : 10-255411

(71)Applicant : TOSHIBA CORP

(22)Date of filing : 09.09.1998

(72)Inventor : MAEDA SEIJI  
KITSU TOSHIKI  
TORII OSAMU  
KANAI TATSUNORI  
YOKOGAWA TAKESHI  
YAO HIROSHI  
TANAKA HISAKO

(30)Priority

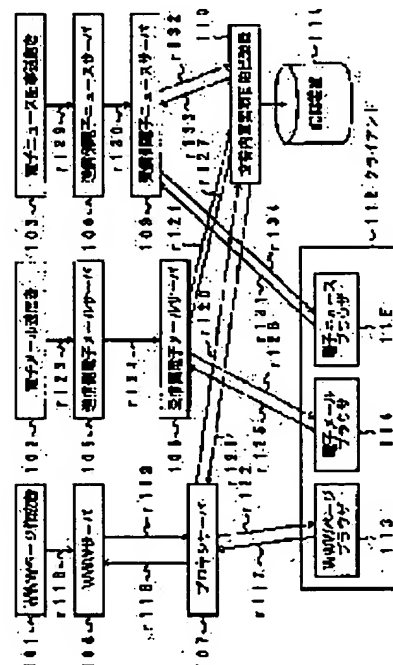
Priority number : 10199584 Priority date : 30.06.1998 Priority country : JP

## (54) STRUCTURED DOCUMENT SAVING METHOD, DATA REPEATING DEVICE, AND DATA TRANSMITTING AND RECEIVING DEVICE

(57)Abstract:

PROBLEM TO BE SOLVED: To automatically save a part including important contents of a structured document without placing any burden on a user by extracting and saving the important part from the structured document by making use of the structure of the structured document obtained on the basis of a previously specified specific extraction standard.

SOLUTION: A WWW page browser 113 having received an indication from the user informs a proxy server 107 of a request to send a WWW page that the user specifies. The proxy server 107 similarly informs a WWW server 104 of the request to send the WWW page and also inputs the received WWW page to a document important item extracting device 110. The document important item extracting device 110 analyzes the structure of the WWW page as a structured document and retrieves whether or not the page contains the important item that the user has specified. When the important item is not found, the important item is extracted and saved in a recording device 111, and a tag indicating that the important item is extracted is added to the WWW page.



## LEGAL STATUS

[Date of request for examination]

28.08.2001

[Date of sending the examiner's decision of rejection]

[Kind of final disposal of application other than the examiner's decision of rejection or application converted registration]

[Date of final disposal for application]

[Patent number] 3715444

[Date of registration] 02.09.2005

[Number of appeal against examiner's decision of rejection]

[Date of requesting appeal against examiner's decision of rejection]

[Date of extinction of right]

Copyright (C); 1998,2003 Japan Patent Office

(11)特許出願公開番号  
特開2000-155756  
(P2000-155756A)

(43)公開日 平成12年6月6日(2000.6.6)

(51)Int.Cl. <sup>7</sup>	識別記号	F I	マークト <sup>(参考)</sup>
G 0 6 F 17/27		C 0 6 F 15/20	5 5 0 F 5 B 0 0 9
17/30		15/40	3 1 0 F 5 B 0 7 5
			3 4 0
			3 7 0 A

審査請求 未請求 請求項の数24 O L (全 64 頁)

(21)出願番号 特願平10-255411

(22)出願日 平成10年9月9日(1998.9.9)

(31)優先權主張番号 特願平10-199584

(32)優先日 平成10年6月30日(1998.6.30)

(33)優先権主張国 日本 (JP)

(71)出願人 000003078

株式会社東芝

神奈川県川崎市幸区堀川町72番地

(72) 発明者 前田 誠司

神奈川県川崎市幸区小向東芝町1番地 株  
式会社東芝研究開発センター内

(72) 発明者 岐津 俊樹

神奈川県川崎市幸区小向東芝町1番地 株  
式会社東芝研究開発センター内

(74) 代理人 100058479

弁理士 鈴江 武彦 (外6名)

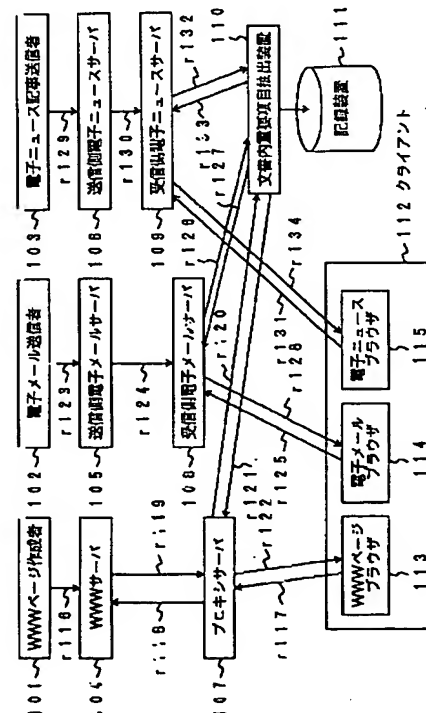
最終頁に続く

(54) 【発明の名称】 構造化文書保存方法並びにデータ中継装置及びデータ送受信装置

(57) 【要約】

【課題】 ユーザに負担をかけずに構造化文書のうち重要な内容を含む部分を自動的に保存するための構造化文書保存方法を提供すること。

【解決手段】 要求された構造化文書を提供するサーバ装置から取得した構造化文書を処理するユーザシステムのために該取得した構造化文書を部分的に自動保存するための構造化文書保存方法であって、ユーザシステムが要求した構造化文書を取得し、予め指定された所定の抽出基準に基づき前記取得した構造化文書から重要部分を抽出して保存するとともに、該取得した構造化文書を前記ユーザシステムに渡すことを特徴とする。例えば、構造化文書から指定されたタグに該当する構造化部分を抜き出して保存する。



# 【特許請求の範囲】

【請求項1】要求された構造化文書を提供するサーバ装置から取得した構造化文書処理するユーザシステムのために該取得した構造化文書を部分的に自動保存するための構造化文書保存方法であって、

ユーザシステムの要求に応じてサーバ装置により提供された構造化文書を取得し、

予め指定された所定の抽出基準に基づき前記取得した構造化文書の構造を利用して該構造化文書から重要部分を抽出して保存するとともに、該取得した構造化文書を前記ユーザシステムに渡すことを特徴とする構造化文書保存方法。

【請求項2】取得した前記構造化文書から重要部分を抽出する際、構造化されている単位での抽出を行い、抽出された前記構造化文書の重要部分を保存する際、複数の構造化文書から別々に抽出された複数の重要部分を内部構造として持たせるようにした構造化されたデータベースに保存することを特徴とする請求項1に記載の構造化文書保存方法。

【請求項3】同一の文書から抽出された複数の重要部分を前記データベースに保存する際、該同一の文書から抽出された複数の重要部分の相互間の関連を示すリンク情報を作成して保存することを特徴とする請求項2に記載の構造化文書保存方法。

【請求項4】前記取得した構造化文書を前記ユーザシステムに渡す際、該構造化文書における重要部分として抽出された部分に、その部分が重要部分として抽出されたことをユーザに提示するための所定の情報を付加した上で、該構造化文書を前記ユーザシステムに渡すことを特徴とする請求項1ないし3のいずれか1項に記載の構造化文書保存方法。

【請求項5】構造化文書から重要部分を抽出するための前記所定の抽出基準には、少なくとも重要部分として抽出する部分の持つべきタグおよび／または単語の指定を含み、

対象となる構造化文書について参照すべき文書型定義情報が指定されている場合には、該対象となる構造化文書からの重要部分の抽出に用いるタグおよび／または単語を、該指定された文書型定義情報に含まれるタグおよび／または単語とすることを特徴とする請求項1ないし4のいずれか1項に記載の構造化文書保存方法。

【請求項6】構造化文書から重要部分を抽出するための前記所定の抽出基準には、少なくとも重要部分として抽出する部分の持つべきタグおよび／または単語の指定を含み、

前記指定されたタグおよび／または単語に関して所定の類語データベースを検索して得たタグおよび／または単語に対する類語をも用いて、前記構造化文書から重要部分を抽出することを特徴とする請求項1ないし5のいずれか1項に記載の構造化文書保存方法。

【請求項7】前記ユーザシステムから、構造化文書から重要部分を抽出するための前記所定の抽出基準の内容についての指定、追加、変更または削除の要求を受け、該要求に従って前記所定の抽出基準の内容を新規登録、追加、変更または削除することを特徴とする請求項1ないし6のいずれか1項に記載の構造化文書保存方法。

【請求項8】前記所定の抽出基準に基づいて重要部分が抽出され前記構造化文書を受け取った前記ユーザシステムにおいて該構造化文書を表示した際にユーザから該所定の抽出基準の内容を変更するための指示が入力された場合に該指示に従って、前記抽出に用いた所定の抽出基準の内容を変更することを特徴とする請求項1ないし7のいずれか1項に記載の構造化文書保存方法。

【請求項9】前記取得した構造化文書から重要部分を抽出して保存する際、予め指定された所定の生成基準に基づいて、抽出された前記重要部分に関連のある付加情報を生成し、抽出された該重要部分と生成された該付加情報とを対応付けて保存することを特徴とする請求項1ないし8のいずれか1項に記載の構造化文書保存方法。

【請求項10】取得した前記構造化文書から重要部分を抽出する際に構造化されている単位での抽出を行うとともに、前記重要部分に関連のある付加情報を構造化されたデータとして生成し、

抽出された前記重要部分と生成された前記付加情報とを保存する際、複数の構造化文書から別々に抽出された複数の重要部分および付加情報を内部構造として持たせるようにした構造化されたデータベースに保存することを特徴とする請求項9に記載の構造化文書保存方法。

【請求項11】要求された構造化文書を提供するサーバ装置から取得した構造化文書処理するユーザシステムのために該取得した構造化文書を部分的に自動保存するための構造化文書保存方法であって、

ユーザシステムの要求に応じてサーバ装置により提供された構造化文書を取得し、

予め指定された所定の抽出基準に基づき前記取得した構造化文書の構造を利用して該構造化文書から重要部分を抽出し、該重要部分に前記要求を出したユーザの識別情報を対応付けて保存するとともに、該取得した構造化文書を前記ユーザシステムに渡すことを特徴とする構造化文書保存方法。

【請求項12】保存されているある重要部分に対してあるユーザからのアクセスが発生した場合に、該重要部分に対応して保存されているアクセス制御情報に基づいて該重要部分を該ユーザがアクセス可能か否かを判断し、アクセス可能であると判断された場合に、該重要部分を該ユーザに送信することを特徴とする請求項11に記載の構造化文書保存方法。

【請求項13】前記重要部分は、構造化された単位で抽出されて、複数の構造化文書から別々に抽出された複数の重要部分を内部構造として持たせるようにした構造化



されたデータベースに保存されたものであり、前記アクセス制御情報は、各重要部分ごとに作成された該重要部分の所有者となるユーザの識別情報および該重要部分に対するアクセスを許可されたユーザの識別情報のリストを記述した内容を持つ構造化されたデータであって、各重要部分に対応付けて前記構造化されたデータベースに保存されたものであることを特徴とする請求項12に記載の構造化文書保存方法。

【請求項14】あるアクセス制御情報に対する内容変更の要求をあるユーザから受けた場合、そのアクセス制御情報に基づいて該ユーザに該アクセス制御情報の内容を変更する権限があるか否かを判断し、変更する権限があると判断された場合に、該アクセス制御情報を要求された内容に従って変更することを特徴とする請求項12または13に記載の構造化文書保存方法。

【請求項15】前記取得した構造化文書から重要部分が抽出された場合に、既に同一の構造化文書から同一の重要部分が抽出され保存されているならば、該取得した構造化文書から抽出された重要部分を新たに保存することはせずに、該既に保存されている同一の重要部分に対応するアクセス制御情報に対して、前記要求を出したユーザが少なくとも該同一の重要部分をアクセス可能となるように、該ユーザの識別情報を付加することを特徴とする請求項11ないし14のいずれか1項に記載の構造化文書保存方法。

【請求項16】前記取得した構造化文書から重要部分が抽出された場合に、既に同一の構造化文書から同一の重要部分が抽出され保存されており、かつ、該既に保存されている同一の重要部分に対応するアクセス制御情報によって既に前記要求を出したユーザが少なくとも該同一の重要部分をアクセス可能となっているならば、該取得した構造化文書から抽出された重要部分の新たな保存も該アクセス制御情報に対する前記要求を出したユーザの識別情報の付加もせずに、該取得した構造化文書を前記ユーザシステムに渡すことを特徴とする請求項15に記載の構造化文書保存方法。

【請求項17】前記重要部分は、構造化された単位で抽出されて、複数の構造化文書から別々に抽出された複数の重要部分を内部構造として持たせるようにした構造化されたデータベースに保存されたものであり、前記アクセス制御情報は、各重要部分ごとに作成された該重要部分の所有者となるユーザの識別情報および該重要部分に対するアクセスを許可されたユーザの識別情報のリストを記述した内容を持つ構造化されたデータであって、各重要部分に対応付けて前記構造化されたデータベースに保存されたものであり、前記既に保存されている同一の重要部分に対応するアクセス制御情報に対して前記要求を出したユーザの識別情報を付加する際、該既に保存されている同一の重要部分に対応するアクセス制御情報に該要求を出したユーザの

識別情報を、少なくとも前記アクセスを許可されたユーザの識別情報として付加することを特徴とする請求項15または16に記載の構造化文書保存方法。

【請求項18】前記重要部分は、構造化された単位で抽出されて、複数の構造化文書から別々に抽出された複数の重要部分を内部構造として持たせるようにした構造化されたデータベースに保存されたものであり、

前記アクセス制御情報は、各重要部分ごとに作成された、該重要部分の所有者となるユーザの識別情報および該重要部分に対するアクセスを許可されたユーザの識別情報のリストを記述した内容を持つ構造化されたデータであって、各重要部分に対応付けて前記構造化されたデータベースに保存されたものであり、

前記既に保存されている同一の重要部分に対応するアクセス制御情報に対して前記要求を出したユーザの識別情報を付加する際、該既に保存されている同一の重要部分に対応するアクセス制御情報とは別に、該要求を出したユーザの識別情報を少なくとも前記所有者となるユーザの識別情報として記述した新たなアクセス制御情報を作成して、これを該既に保存されている同一の重要部分に対応付けて保存することを特徴とする請求項15または16に記載の構造化文書保存方法。

【請求項19】前記重要部分は、構造化された単位で抽出されて、複数の構造化文書から別々に抽出された複数の重要部分を内部構造として持たせるようにした構造化されたデータベースに保存されたものであり、

前記アクセス制御情報は、各重要部分ごとに作成された該重要部分にアクセス可能なユーザの識別情報のリストを記述した内容を持つ構造化されたデータであって、各重要部分に対応付けて前記構造化されたデータベースに保存されたものであり、

前記既に保存されている同一の重要部分に対応するアクセス制御情報に対して前記要求を出したユーザの識別情報を付加する際、該既に保存されている同一の重要部分に対応するアクセス制御情報に該要求を出したユーザの識別情報を付加することを特徴とする請求項15または16に記載の構造化文書保存方法。

【請求項20】保存されているある重要部分に対する編集要求があった場合に、該重要部分に対応するアクセス制御情報の内容が複数のユーザの該重要部分に対するアクセスを可能とする状態になっているならば、該重要部分のデータを複写して別途同一の内容を持つ重要部分を生成するとともに、該編集要求を出したユーザが該生成された重要部分に対して編集可能となるような形で該ユーザの識別情報を含んだアクセス制御情報を作成して、これら生成された重要部分と作成されたアクセス制御情報とを対応付けて保存することを特徴とする請求項15ないし19のいずれか1項に記載の構造化文書保存方法。

【請求項21】データ送受信装置が要求した構造化文書

を通信路を介して該当するサーバ装置から取得する手段と、

予め指定された所定の抽出基準に基づき前記取得した構造化文書の構造を利用して該構造化文書から重要部分を抽出して保存する手段と、

前記取得した構造化文書を前記データ送受信装置に転送する手段とを備えたことを特徴とするデータ中継装置。

【請求項22】要求された構造化文書を通信路を介して該当するサーバ装置から取得する手段と、

予め指定された所定の抽出基準に基づき前記取得した構造化文書の構造を利用して該構造化文書から重要部分を抽出して保存する手段と、

前記取得した構造化文書を表示する手段とを備えたことを特徴とするデータ送受信装置。

【請求項23】データ送受信装置が要求した構造化文書を通信路を介して該当するサーバ装置から取得する手段と、

予め指定された所定の抽出基準に基づき前記取得した構造化文書の構造を利用して該構造化文書から重要部分を抽出し、該重要部分に前記要求を出したユーザの識別情報を対応付けて保存する手段と、

前記取得した構造化文書を前記データ送受信装置に転送する手段とを備えたことを特徴とするデータ中継装置。

【請求項24】要求された構造化文書を通信路を介して該当するサーバ装置から取得する手段と、

予め指定された所定の抽出基準に基づき前記取得した構造化文書の構造を利用して該構造化文書から重要部分を抽出し、該重要部分に前記要求を出したユーザの識別情報を対応付けて保存する手段と、

前記取得した構造化文書を表示する手段とを備えたことを特徴とするデータ送受信装置。

【発明の詳細な説明】

【0001】

【発明の属する技術分野】本発明は、WWWに蓄積されている文書や電子メール等の構造化文書を交換するシステムに用いられ、その構造化文書を自動保存する機能を有するデータ中継装置及びデータ送受信装置並びに構造化文書保存方法に関する。

【0002】

【従来の技術】構造化文書とは、文書中に組み込まれたタグと呼ばれる区切りなどを用いて、文章の構造を表現することができる文書である。また、構造化文書の中にはDTD (Document Type Definition) と呼ばれる文書構造を規定する方法を提供するものもある。このような文書としては、SGML文書、XML文書、HTML文書や、電子メールあるいは電子ニュースで伝送される構造化文書等がある。このような構造化文書では、その文書の構造を容易に解析することができる。

【0003】World Wide Web (ワールド

ワイドウェブ；以下、WWWと略す)、電子メール、電子ニュース等では、インターネットやイントラネットなどの通信路を介して構造化文書を交換する。

【0004】以下、WWWシステムについて説明する。

【0005】WWWでは、構造化文書配信サーバであるWWWサーバおよび構造化文書表示装置であるブラウザを通信路を介して接続し、構造化文書を交換する。構造化文書を参照する場合には、ブラウザはWWWサーバに対し必要としている構造化文書の送信を要求し、WWWサーバはその応答として構造化文書を送信する。WWWで交換される構造化文書は、WWWページと呼ばれることもある。

【0006】通常のWWWでは、通信路のデータ伝送量の増大とWWWサーバへの負荷の集中を抑えるために、WWWサーバとブラウザとの間にプロキシサーバと呼ばれるデータ中継装置を置くことが多い。このプロキシサーバは、ブラウザとWWWサーバの間に位置し、ブラウザからの要求をWWWサーバに中継する機能を持つ。

【0007】プロキシサーバは、ブラウザから構造化文書の送信要求を受信し、代わりにWWWサーバに構造化文書の送信要求を送信する。次に、WWWサーバより送信された構造化文書をブラウザに転送する。プロキシサーバは、WWWサーバとブラウザとの間に1台だけ配置しても良いし、複数配置しても良い。

【0008】このようなプロキシサーバの中には転送する複数の構造化文書を、プロキシサーバ内に一時的に保存する機能を持つものがある。このような複数の構造化文書を一時的に保存する機能を持つデータ転送装置のことをキャッシュサーバと呼ぶこともある。

【0009】複数の構造化文書を一時的に保存するプロキシサーバでは、ブラウザから送信要求があった構造化文書を一時的に保存しておく。同じ構造化文書に対してブラウザより送信要求があった場合には、再度、WWWサーバに送信要求を転送するのではなく、代わりに一時的に保存しておいた構造化文書をブラウザに送信する。

【0010】その際に、一時的に保存してある構造化文書が最新であるかWWWサーバに確認する機能を持つプロキシサーバも存在する。

【0011】通常、プロキシサーバでは、ある決められた一定量だけ複数の構造化文書を保存する。一時的に構造化文書を保存するに際しては、その構造化文書が、どのブラウザからの要求によるものであるかを区別しない。一定量を越える構造化文書を保存しようとした場合には、ある規則に従い、一時的に保存してある構造化文書を削除する。この際に用いられる規則としては、構造化文書のうち最後に参照された時刻の古い順に構造化文書を削除するというような単純な規則が良く用いられている。構造化文書を一時的に保存する記憶装置としては、例えば磁気ディスク装置などが用いられている。

【0012】ブラウザなどの表示装置においても、一定

量の複数の構造化文書を一時的に保存する機能を持つものがある。このようなブラウザでは、ユーザが要求した構造化文書がブラウザ内に一時的に保存してある構造化文書である場合には、再度、WWWサーバに送信要求を転送するのではなく、代わりに一時的に保存しておいた構造化文書を表示する。

【0013】その際に、一時的に保存しておいた構造化文書が最新であるかWWWサーバやプロキシサーバに確認する機能を持つものもある。

【0014】通常、このようなブラウザでは、ある決められた一定量だけ複数の構造化文書を保存する。その一定量を越える構造化文書を保存しようとした場合には、ある規則に従い、一時的に保存してある構造化文書を削除する。この際に用いられる規則としては、例えば構造化文書のうち最後に参照された時刻の古い順に削除するというような単純な規則が良く用いられている。構造化文書を一時的に保存する記憶装置としては、例えば磁気ディスク装置などが用いられている。

【0015】次に、電子メールシステムについて説明する。

【0016】電子メールシステムは、電子メールサーバと電子メール送受信装置などで構成され、それぞれはインターネットやイントラネットなどの通信路を介して接続される。

【0017】電子メールの送信者は、電子メール送受信装置を用いて電子メールを作成し、受信者のアドレスを指定して電子メールを送信する。電子メールを受信した電子メールサーバは、送信者が指定した受信者のアドレスにしたがって、宛先の電子メールサーバに、または直接送信できない場合には、該電子メールを中継する電子メールサーバを検索し、電子メールを転送する。

【0018】宛先の電子メールサーバは、転送されてきた電子メールを宛先ユーザ毎に分類して格納する。電子メールの受信者は、電子メール送受信装置を用いて電子メールサーバより、受信された電子メールを受信し、電子メールを閲覧する。最終的に電子メールは電子メール格納サーバあるいは電子メール送受信装置あるいはそれら両方に格納される。このような電子メールの文書として構造化文書が用いられる場合がある。

【0019】通常、電子メールは、電子メールサーバに附属の電子メール保存装置の容量分だけ格納することができ、それを越える場合には電子メールサーバは電子メールを受信することが出来ない。この場合には、通常はユーザが明示的に該ユーザ宛の電子メールを削除する。

【0020】次に、電子ニュースシステムについて説明する。

【0021】電子ニュースシステムは、電子ニュースサーバと電子ニュース送受信装置などで構成され、それぞれはインターネットやイントラネットなどの通信路を介して接続される。

【0022】電子ニュース記事の投稿者は、電子ニュース送受信装置を用いて電子ニュース記事を作成し、電子ニュースサーバに送信する。電子ニュース記事を受信した電子ニュースサーバは、パケツリレー式に通信路で接続されている電子ニュースサーバに電子ニュース記事を送信する。なお、電子ニュースサーバは、受信した電子ニュース記事の一覧を保持しており、一度受信した電子ニュース記事を重複して受信しないようになっている。

【0023】電子ニュース記事の購読者は、電子ニュース送受信装置を用いて電子ニュースサーバから電子ニュース記事を受信する。

【0024】通常、電子ニュース記事は、電子ニュースサーバに附属の電子ニュース記事保存装置の容量分だけ格納することができ、それを越える場合には電子ニュースサーバは電子ニュース記事を受信することが出来ない。この場合には、電子ニュースサーバ上で稼働している電子ニュースシステムが自動的に、もしくは電子ニュースサーバの管理者が明示的に、電子ニュース記事を削除する。

【0025】

【発明が解決しようとする課題】従来、WWWは、

(1) 構造化文書配信サーバであるWWWサーバおよび構造化文書表示装置であるブラウザを通信路を介して接続し、構造化文書を交換する形態

(2) 構造化文書配信サーバであるWWWサーバと構造化文書表示装置であるブラウザの間に、構造化文書の要求と応答を中継する1つまたは複数のプロキシサーバを配置し、これらを介して接続し、構造化文書を交換する形態

(3) 構造化文書配信サーバであるWWWサーバと構造化文書表示装置であるブラウザの間に、構造化文書を一時的に保存する1つまたは複数のプロキシサーバを配置し、これらを介して接続し、構造化文書を交換する形態であった。

【0026】しかし、上記の(1)のようにWWWサーバとブラウザの間にプロキシサーバが存在しない場合には、直接構造化文書が送信されてくるため、ユーザはそれを受信した時点で、重要な項目を含む文書かどうかを判断し、必要に応じてその文書を保存しなければならないという問題があった。

【0027】また、上記の(2)のようにWWWサーバとブラウザの間にプロキシサーバが存在する場合でも、プロキシサーバが単に要求と応答を中継する機能しか持たない場合には、ブラウザに直接構造化文書が送信されてくるため、ユーザはそれを受信した時点で、重要な項目を含む文書かどうかを判断し、必要に応じてその文書を保存しなければならないという問題があった。

【0028】また、上記の(3)のようにWWWサーバとブラウザの間に構造化文書を一時的に保存する機能を有するプロキシサーバが配置されている場合やブラウザ

自身が受信した構造化文書を一時的に保存する機能を有する場合、プロキシサーバやブラウザなどで一時的に保存された構造化文書は、ある規則に従い削除される。この際に用いられる規則としては、その構造化文書が参照された時刻の古い順に削除するというような単純な規則が一般的である。そのため、構造化文書内に重要な内容が含まれていた場合においても、その文書が重要であるかどうかを判断せずに削除してしまうので、後に重要な内容を含む構造化文書を参照することができないという問題があった。

【0029】一方、重要な内容を含む構造化文書がすぐに削除されることを防ぐために、より多くの構造化文書を保存する方法もあるが、この場合、保存するために多くの容量を持った記憶装置を必要とするという問題があった。

【0030】さらに、この場合、保存した構造化文書が多いために、重要な内容を含む構造化文書を参照または検索する時に非常に時間がかかるという問題があった。

【0031】また同様に、電子メールに含まれる構造化文書も電子メールサーバもしくは電子メール送受信装置に接続されている記憶装置の容量分だけ保存することができるが、それを超える場合には電子メールを受信することが出来ない。そのため、ユーザが明示的に電子メールを削除しなければならないが、その場合、ユーザがその構造化文書の重要性を確認しながら削除を行わなければならないという問題があった。

【0032】この場合も、重要な内容を含む構造化文書がすぐに削除されることを防ぐために、より多くの構造化文書を保存する方法もあるが、この場合、保存するために多くの容量を持った記憶装置を必要とするという問題があった。さらに、この場合、保存した構造化文書が多いために、重要な内容を含む構造化文書を参照または検索する時に非常に時間がかかるという問題があった。

【0033】また同様に、電子ニュース記事に含まれる構造化文書も、電子ニュースサーバに附属の電子ニュース記事保存装置の容量分だけ格納することができるが、それを超える場合には電子ニュース記事を受信することが出来ない。そのため、電子ニュースサーバ上で稼働している電子ニュースシステムが自動的に、もしくは電子ニュースサーバの管理者が明示的に電子ニュース記事を削除しなければならないが、電子ニュース記事に含まれる構造化文書内に重要な内容が含まれていた場合においても、その文書が重要であるかどうかを判断せずに削除してしまうので、後に重要な内容を含む構造化文書を参照することができないという問題があった。

【0034】この場合も、重要な内容を含む構造化文書がすぐに削除されることを防ぐために、より多くの構造化文書を保存する方法もあるが、この場合、保存するために多くの容量を持った記憶装置を必要とするという問題があった。さらに、この場合、保存した構造化文書が

多いために、重要な内容を含む構造化文書を参照または検索する時に非常に時間がかかるという問題があった。

【0035】本発明は、上記事情を考慮してなされたもので、ユーザに負担をかけずに構造化文書のうち重要な内容を含む部分を自動的に保存するための構造化文書保存方法並びにデータ中継装置及びデータ送受信装置を提供することを目的とする。

【0036】

【課題を解決するための手段】本発明（請求項1）は、要求された構造化文書を提供するサーバ装置から取得した構造化文書を処理するユーザシステムのために該取得した構造化文書を部分的に自動保存するための構造化文書保存方法であって、ユーザシステムの要求に応じてサーバ装置により提供された構造化文書を取得し、予め指定された所定の抽出基準に基づき前記取得した構造化文書の構造を利用して該構造化文書から重要部分（重要項目）を抽出して保存するとともに、該取得した構造化文書を前記ユーザシステムに渡すことを特徴とする。

【0037】構造化文書は、例えば、SGML文書、XML文書、HTML文書や、電子メールあるいは電子ニュースで伝送される構造化文書などである。

【0038】サーバ装置は、例えば、WWWサーバ、電子メールサーバ、電子ニュースサーバなどである。

【0039】例えば、本発明の構造化文書保存方法に係るモジュール（ここでは構造化文書保存モジュールと呼ぶ）をプロキシサーバなどのデータ中継装置に一機能として搭載して実現する場合、構造化文書の要求や構造化文書の配信の流れは例えば次のようになる。なお、この場合、ユーザシステムは、例えば、WWWページブラウザあるいは電子メールブラウザあるいは電子ニュースブラウザなどが起動するユーザ端末もしくはクライアントシステムなどである。ユーザ端末システム→（取得要求転送）→プロキシサーバ→（取得要求転送）→WWWサーバ→（構造化文書配信）→プロキシサーバ（重要部分の抽出・保存）→（構造化文書配信）→ユーザ端末システム（取得した構造化文書の表示等）また、例えば、構造化文書保存モジュールをユーザ端末もしくはクライアントシステムなどのデータ送受信装置に一機能として搭載する場合、構造化文書の要求や構造化文書の配信の流れは例えば次のようになる。なお、この場合、ユーザシステムは、例えば、データ送受信装置上で起動するWWWページブラウザあるいは電子メールブラウザあるいは電子ニュースブラウザなどのアプリケーションもしくはプロセスなどであり、構造化文書保存モジュールも同じデータ送受信装置上で起動するアプリケーションもしくはプロセスなどとなる。ユーザ端末上で起動しているブラウザなどのアプリケーション→（取得要求）→ユーザ端末上で起動している構造化文書保存モジュールに相当するアプリケーション→（取得要求転送）→WWWサーバ→（構造化文書転送）→ユーザ端末上で起動している

構造化文書保存モジュールに相当するアプリケーション  
 {重要部分の抽出・保存} → (構造化文書) → ユーザ端  
 末上で起動しているブラウザなどのアプリケーション  
 {取得した構造化文書の表示など}

この場合、取得要求のメッセージの転送や、構造化文書をユーザシステムに渡すための通信は、ユーザ端末等の内部での通信(例えばプロセス間の通信)となる。端末内部の通信は、実際にメッセージや文書などのデータの受け渡しを行う代わりに、当該データが記憶されている位置を示す情報を受け渡しするようにしてもよい。

【0040】なお、構造化文書保存モジュールを独立した装置として実現してもよい。

【0041】重要部分を抽出するための所定の抽出基準は、例えば、ユーザから抽出すべき項目の分類ごとにタグの指定および/または単語の指定として与えられる。そして、例えば、取得した構造化文書からユーザ指定のタグに一致する部分やユーザ指定の単語を含む部分を抽出し、その部分を文書構造を維持して保存する(例えば当該開始タグと終了タグに挟まれた部分を保存する)。

【0042】本発明では、まず、ユーザが例えばブラウザなどのアプリケーション経由で所望の構造化文書をサーバ装置に要求する。その構造化文書がサーバ装置から送られてくる途中で、構造化文書保存モジュールは、その構造化文書を取り上げ、アプリケーションにかかわらずその中身を見て、重要部分を抽出し所定の記憶装置に保存する。

【0043】構造化文書は、重要部分の抽出が行われてから(もしくは構造化文書保存モジュールが構造化文書を読み込んでから)、ユーザ(ブラウザ等)に渡される。

【0044】ユーザ(ブラウザ等)は、サーバ装置から転送されてきたものとして、該構造化文書を受け取る。そして、例えば、その表示などを行う。

【0045】本発明によれば、ユーザに代わって、受信した構造化文書内からユーザにとって重要である項目を自動的に抽出し記憶装置に保存することができる。構造化文書内の重要部分の抽出・保存作業は自動的に行われるので、ユーザは、受信した構造化文書ごとに自ら重要な項目を抽出し保存する煩わしさから開放され、また抽出し忘れることも保存し忘れることも防ぐことができる。

【0046】なお、ユーザがある項目についての抽出を中止した場合(すなわち、そのように抽出基準を変更した場合)でも、そのときまでに抽出された重要部分は、削除されずに保存されているため、後に再び、該重要部分が必要になった場合でも、該重要部分の読み出し、検索などの操作が可能である。

【0047】本発明においては、好ましくは、取得した前記構造化文書から重要部分を抽出する際、構造化されている単位での抽出を行い、抽出された前記構造化文書

の重要部分を保存する際、複数の構造化文書から別々に抽出された複数の重要部分を内部構造として持たせるようにした構造化されたデータベースに保存するようにしてもよい。すなわち、例えば保存する際に、文書という単位は残さず、複数の文書から抽出した重要部分のみがデータベースになるようにしてもよい。

【0048】また、上記の場合に、好ましくは、同一の文書から抽出された複数の重要部分を前記データベースに保存する際、該同一の文書から抽出された複数の重要部分の相互間の関連を示すリンク情報を作成して保存するようにしてもよい。すなわち、保存する際に文書という単位を残さない場合に、1つの文書から抽出した複数の重要部分間の関連が分からなくなならないように、リンクして保存すると好ましい。これによって、どの重要部分とどの重要部分がもともと同一の文書にあったものであるかを容易に知ることができる。

【0049】また、好ましくは、前記取得した構造化文書を前記ユーザシステムに渡す際、該構造化文書における重要部分として抽出された部分に、その部分が重要部分として抽出されたことをユーザに提示するための所定の情報を付加した上で、該構造化文書を前記ユーザシステムに渡すようにしてもよい。所定の情報は、例えば、所定のタグである。例えば、反転、強調、網かけなどを指示するタグを付加しておけば、ユーザ側で該構造化文書を表示した際に、重要部分として抽出された部分を他の部分とは区別して表示させることができるので、ユーザは、重要部分として抽出された部分を容易に把握することができる。また、ユーザは、抽出された重要部分を確認することができ、ユーザの意図通りに重要部分の抽出・保存が行われているかどうかを確認することができる。

【0050】また、好ましくは、構造化文書から重要部分を抽出するための前記所定の抽出基準には、少なくとも重要部分として抽出する部分の持つべきタグおよび/または単語の指定を含み、対象となる構造化文書について参照すべき文書型定義情報(例えば、Document Type Definition; DTD)が指定されている場合には、該対象となる構造化文書からの重要部分の抽出に用いるタグおよび/または単語を、該指定された文書型定義情報に含まれるタグおよび/または単語とするようにしてもよい。

【0051】DTDは、ある構造化文書の中で、どのようなタグを用いることが出来るか、それらのタグの階層関係はどのようになっているかなどを定義し、構造化文書の論理構造を明確に定義するものである。構造化文書の抽出タグ名を指定する際に、抽出する構造化文書の参照するDTDが指定されている場合には、DTDによって構造化文書内に含まれるタグ名の種類が決まるので、構造化文書内に含まれる可能性のあるタグのみを抽出タグ名に指定できるようにすると好ましい。これによ



て、抽出すべきものとして指定するタグ名の設定の誤りを防ぐことができる。

【0052】また、好ましくは、構造化文書から重要部分を抽出するための前記所定の抽出基準には、少なくとも重要部分として抽出する部分の持つべきタグおよび／または単語の指定を含み、前記指定されたタグおよび／または単語に関して所定の類語データベースを検索して得たタグおよび／または単語に対する類語をも用いて、前記構造化文書から重要部分を抽出するようにしてもよい。このように類語を用いるのを前記所定の抽出基準として類語を用いることが指定されている場合とすることもできる。これによって、重要部分の取りこぼしを防ぐなど、より効果的な抽出・保存ができる。

【0053】また、好ましくは、前記ユーザシステムから、構造化文書から重要部分を抽出するための前記所定の抽出基準の内容についての指定、追加、変更または削除の要求を受け、該要求に従って前記所定の抽出基準の内容を新規登録、追加、変更または削除するようにしてもよい。

【0054】本発明によれば、ユーザは、抽出・保存させたい重要な項目に関する条件などの指定、追加、変更、削除を行うことができ、ユーザは必要に応じて、適切な重要な項目を自動的に抽出・保存させることができる。

【0055】また、好ましくは、前記所定の抽出基準に基づいて重要部分が抽出され前記構造化文書を受け取った前記ユーザシステムにおいて該構造化文書を表示した際に（該ユーザシステムに対して）ユーザから該所定の抽出基準の内容を変更するための指示が入力された場合に該指示（ユーザシステムを介して構造化文書保存モジュールに与えられる）に従って、前記抽出に用いた所定の抽出基準の内容を変更するようにしてもよい。例えば、ユーザ側で受け取った構造化文書を表示した際に、ユーザは、その抽出結果を見て判断するなどして抽出基準の内容を変更すべきと考えた場合には、変更内容についての指示を入力し、これが当該抽出基準の内容に反映される。

【0056】また、好ましくは、前記取得した構造化文書から重要部分を抽出して保存する際、予め指定された所定の生成基準に基づいて、抽出された前記重要部分に関連のある付加情報（付加項目）を生成し、抽出された該重要部分と生成された該付加情報とを対応付けて保存するようにしてもよい。

【0057】本発明によれば、ユーザに代わって、受信した構造化文書内からユーザにとって重要である項目を自動的に抽出し、抽出した項目に関連がある付加情報（例えば、構造化文書の重要情報を抽出した日時、構造化文書が存在する位置（例えば、URL）、構造化文書の重要項目以外の項目（すべてでもよいし指定してもよい）が存在する位置など）を生成し、両者を対応付けて

記憶装置に保存することができる。抽出、生成、保存の作業は自動的になされるので、ユーザは、閲覧した構造化文書ごとに自力で重要部分や付加情報を抽出し保存する煩わしさから開放され、抽出し忘れることも保存し忘れることも防ぐことができる。

【0058】また、好ましくは、取得した前記構造化文書から重要部分を抽出する際に構造化されている単位での抽出を行うとともに、前記重要部分に関連のある付加情報を構造化されたデータとして生成し、抽出された前記重要部分と生成された前記付加情報とを保存する際、複数の構造化文書から別々に抽出された複数の重要部分および付加情報を内部構造として持たせるようにした構造化されたデータベースに保存するようにしてもよい。

【0059】なお、本発明においては、抽出・保存された重要部分、または重要部分および付加情報をユーザが検索できる機能を設けることが好ましい。これによって、あらかじめ保存してある構造化文書内の重要部分や付加情報の中から所望の情報を得ることが可能となる。また、例えば、ユーザにとって重要な項目を抽出・保存するための抽出基準の指示の仕方が不十分であったために、ユーザにとって重要である項目が抽出されなかった場合、付加情報として構造化文書が存在する位置を保持していれば、再度構造化文書全体の転送を要求することで、所望の情報を得ることが可能となる。

【0060】なお、以上の請求項1～請求項10に関して説明した上記の各種事項は基本的に請求項11以降の発明についても成り立つものである。

【0061】本発明（請求項11）は、要求された構造化文書を提供するサーバ装置から取得した構造化文書を処理するユーザシステムのために該取得した構造化文書を部分的に自動保存するための構造化文書保存方法であって、ユーザシステムの要求に応じてサーバ装置により提供された構造化文書を取得し、予め指定された所定の抽出基準に基づき前記取得した構造化文書の構造を利用して該構造化文書から重要部分を抽出し、該重要部分に前記要求を出したユーザの識別情報を対応付けて保存するとともに、該取得した構造化文書を前記ユーザシステムに渡すことを特徴とする。

【0062】このためには、抽出した重要部分にユーザ識別情報を付して保存するようにしてもよいし、抽出した重要部分をユーザ毎に分離して保存するようにしてもよい。

【0063】本発明によれば、ユーザが指定した抽出すべき重要な項目を自動的に構造化文書から抽出し保存するが、抽出した重要部分をユーザとの対応を取って保存することにより、特定のユーザが保存するよう指定した重要な項目の読み出し、検索が容易になる。

【0064】好ましくは、保存されているある重要部分に対してあるユーザからのアクセスが発生した場合に、該重要部分に対応して保存されているアクセス制御情報

(例えば、アクセス許可リスト)に基づいて該重要部分を該ユーザがアクセス可能か否かを判断し、アクセス可能であると判断された場合に、該重要部分を該ユーザに送信するようにしてもよい。アクセス制御情報には、例えば、その重要部分をアクセス可能なユーザが記述されている。なお、アクセス可能でないと判断された場合には、該重要部分の該ユーザへの送信を拒否することになる。また、好ましくは、前記重要部分は、構造化された単位で抽出されて、複数の構造化文書から別々に抽出された複数の重要部分を内部構造として持たせるようにした構造化されたデータベースに保存されたものであり、前記アクセス制御情報は、各重要部分ごとに作成された該重要部分の所有者となるユーザの識別情報および該重要部分に対するアクセスを許可されたユーザの識別情報のリストを記述した内容を持つ構造化されたデータであって、各重要部分に対応付けて前記構造化されたデータベースに保存されたものであるようにしてもよい。このように、保存した重要部分毎に、アクセス制御情報を持つことによって、保存した重要部分を複数のユーザ間で、容易に共有することが可能になり、記憶装置の効率的な利用が可能となる。

【0065】また、好ましくは、あるアクセス制御情報に対する内容変更の要求をあるユーザから受けた場合、そのアクセス制御情報に基づいて該ユーザに該アクセス制御情報の内容を変更する権限があるか否かを判断し、変更する権限があると判断された場合に、該アクセス制御情報を要求された内容に従って変更するようにしてもよい。例えば、アクセス制御情報は各重要部分ごとに用意され、アクセス制御情報にはその重要部分の所有者となるユーザ(例えば、その重要部分を抽出するものとなった構造化文書の取得を要求したユーザ)および該所有者以外で該重要部分に対するアクセスを許可されたユーザのリストが記述され、その重要部分の所有者のみが変更の権限を有するものとする。例えば、そのアクセス制御情報に変更を要求したユーザが所有者として登録されている場合には、該ユーザに該アクセス制御情報の内容を変更する権限があると判断し、該アクセス制御情報を要求された内容に従って変更するようにしてもよい。

【0066】なお、アクセス制御情報は、重要項目のデータベースとは別に保存するようにしてもよい。

【0067】また、本発明(請求項15)のように、請求項11~14の発明において、前記取得した構造化文書から重要部分が抽出された場合に、既に同一の構造化文書から同一の重要部分が抽出され保存されているならば、該取得した構造化文書から抽出された重要部分を新たに保存することはせずに、該既に保存されている同一の重要部分に対応するアクセス制御情報に対して、前記要求を出したユーザが少なくとも該同一の重要部分をアクセス可能となるように、該ユーザの識別情報を付加するようにしてもよい。

【0068】この場合に、好ましくは、前記取得した構造化文書から重要部分が抽出された場合に、既に同一の構造化文書から同一の重要部分が抽出され保存されており、かつ、該既に保存されている同一の重要部分に対応するアクセス制御情報によって既に前記要求を出したユーザが少なくとも該同一の重要部分をアクセス可能となっているならば、該取得した構造化文書から抽出された重要部分の新たな保存も該アクセス制御情報に対する前記要求を出したユーザの識別情報の付加もせずに、該取得した構造化文書を前記ユーザシステムに渡すようにしてもよい。

【0069】例えば、複数のユーザが同一の重要項目抽出条件を設定しており(もしくは設定した重要項目抽出条件の少なくとも一部に共通するものがあり)、かつ該複数のユーザが同一の構造化文書を要求した場合に、ユーザ間で共通の重要項目が抽出されることから、受信した構造化文書から抽出された重要項目のうちユーザ間で共通の重要項目についてはこれを共有できる形態で保存することにより、記憶装置のより効率的な利用が可能となる。

【0070】好ましくは、前記重要部分は、構造化された単位で抽出されて、複数の構造化文書から別々に抽出された複数の重要部分を内部構造として持たせるようにした構造化されたデータベースに保存されたものであり、前記アクセス制御情報は、各重要部分ごとに作成された該重要部分の所有者となるユーザの識別情報および該重要部分に対するアクセスを許可されたユーザの識別情報のリストを記述した内容を持つ構造化されたデータであって、各重要部分に対応付けて前記構造化されたデータベースに保存されたものであり、前記既に保存されている同一の重要部分に対応するアクセス制御情報に対して前記要求を出したユーザの識別情報を付加する際、該既に保存されている同一の重要部分に対応するアクセス制御情報に該要求を出したユーザの識別情報を、少なくとも前記アクセスを許可されたユーザの識別情報として(前記アクセスを許可されたユーザの識別情報として、または前記所有者となるユーザの識別情報および前記アクセスを許可されたユーザの識別情報の両方として)付加するようにしてもよい。

【0071】あるいは、好ましくは、前記重要部分は、構造化された単位で抽出されて、複数の構造化文書から別々に抽出された複数の重要部分を内部構造として持たせるようにした構造化されたデータベースに保存されたものであり、前記アクセス制御情報は、各重要部分ごとに作成された、該重要部分の所有者となるユーザの識別情報および該重要部分に対するアクセスを許可されたユーザの識別情報のリストを記述した内容を持つ構造化されたデータであって、各重要部分に対応付けて前記構造化されたデータベースに保存されたものであり、前記既に保存されている同一の重要部分に対応するアクセス制

御情報に対して前記要求を出したユーザの識別情報を付加する際、該既に保存されている同一の重要部分に対応するアクセス制御情報とは別に、該要求を出したユーザの識別情報を少なくとも前記所有者となるユーザの識別情報として（前記所有者となるユーザの識別情報として、または前記所有者となるユーザの識別情報および前記アクセスを許可されたユーザの識別情報の両方として）記述した新たなアクセス制御情報を作成して、これを該既に保存されている同一の重要部分に対応付けて保存するようにしてもよい。

【0072】あるいは、好ましくは、前記重要部分は、構造化された単位で抽出されて、複数の構造化文書から別々に抽出された複数の重要部分を内部構造として持たせるようにした構造化されたデータベースに保存されたものであり、前記アクセス制御情報は、各重要部分ごとに作成された該重要部分にアクセス可能なユーザの識別情報のリストを記述した内容を持つ構造化されたデータであって、各重要部分に対応付けて前記構造化されたデータベースに保存されたものであり、前記既に保存されている同一の重要部分に対応するアクセス制御情報に対して前記要求を出したユーザの識別情報を付加する際、該既に保存されている同一の重要部分に対応するアクセス制御情報に該要求を出したユーザの識別情報を付加するようにしてもよい。

【0073】また、好ましくは、保存されているある重要部分に対する編集要求があった場合に、該重要部分に対応するアクセス制御情報の内容が複数のユーザの該重要部分に対するアクセスを可能とする状態になっているならば、該重要部分のデータを複写して別途同一の内容を持つ重要部分を生成するとともに、該編集要求を出したユーザが該生成された重要部分に対して編集可能となるような形で該ユーザの識別情報を含んだアクセス制御情報を作成して、これら生成された重要部分と作成されたアクセス制御情報とを対応付けて保存するようにしてもよい。

【0074】もちろん、請求項11～請求項20の各々の発明に対しても、請求項2～請求項10の発明を適宜組み合わせ適用することができる。

【0075】また、本発明（請求項21）に係るデータ中継装置（文書内重要項目抽出装置）は、データ送受信装置が要求した構造化文書を通信路を介して該当するサーバ装置から取得する手段と、予め指定された所定の抽出基準に基づき前記取得した構造化文書の構造を利用して該構造化文書から重要部分を抽出して保存する手段と、前記取得した構造化文書を前記データ送受信装置に転送する手段とを備えたことを特徴とする。

【0076】また、本発明（請求項22）に係るデータ送受信装置（文書内重要項目抽出装置）は、要求された構造化文書を通信路を介して該当するサーバ装置から取得する手段と、予め指定された所定の抽出基準に基づき

前記取得した構造化文書の構造を利用して該構造化文書から重要部分を抽出して保存する手段と、前記取得した構造化文書を表示する手段とを備えたことを特徴とする。

【0077】また、本発明（請求項23）に係るデータ中継装置（文書内重要項目抽出装置）は、データ送受信装置が要求した構造化文書を通信路を介して該当するサーバ装置から取得する手段と、予め指定された所定の抽出基準に基づき前記取得した構造化文書の構造を利用して該構造化文書から重要部分を抽出し、該重要部分に前記要求を出したユーザの識別情報を対応付けて保存する手段と、前記取得した構造化文書を前記データ送受信装置に転送する手段とを備えたことを特徴とする。

【0078】また、本発明（請求項24）に係るデータ送受信装置（文書内重要項目抽出装置）は、要求された構造化文書を通信路を介して該当するサーバ装置から取得する手段と、予め指定された所定の抽出基準に基づき前記取得した構造化文書の構造を利用して該構造化文書から重要部分を抽出し、該重要部分に前記要求を出したユーザの識別情報を対応付けて保存する手段と、前記取得した構造化文書を表示する手段とを備えたことを特徴とする。

【0079】なお、装置に係る本発明は方法に係る発明としても成立し、方法に係る本発明は装置に係る発明としても成立する。

【0080】また、装置または方法に係る本発明は、コンピュータに当該発明に相当する手順を実行させるための（あるいはコンピュータを当該発明に相当する手段として機能させるための、あるいはコンピュータに当該発明に相当する機能を実現させるための）プログラムを記録したコンピュータ読取り可能な記録媒体としても成立する。

【0081】

【発明の実施の形態】以下、図面を参照しながら発明の実施の形態を説明する。

【0082】本発明における重要項目抽出装置は、基本的な機能として、重要項目を指定する機能と、通信路を介して送信されてくる構造化文書から、該重要項目に基づいて自動的に重要項目を抽出し、保存する機能とを提供するものである。また、保存した重要項目に対する検索機能をも提供することができるものである。また、保存した重要項目を、複数のユーザで共有する機能をも提供することができるものである。さらに、重要項目に関連する付加項目を自動的に生成し、抽出した重要項目と組にして保存する機能をも提供することができるものであり、この場合に、さらに検索機能や共有機能をも提供することができるものである。

【0083】ここで、構造化文書とは、タグと呼ばれる文書の構造を表現するためのテキストとタグ以外のデータとが混在する文書のことである。構造化文書には、例



例えば、SGML文書、XML文書、HTML文書や、電子メールあるいは電子ニュースで伝送される構造化文書等がある。

【0084】構造化文書は、タグによって、その論理的構造が決定される。各タグは、特定の符号で開始し、特定の符号で終了する。タグには、開始タグ、終了タグ、空タグなどが存在する。また、各タグは、タグ名や属性名、属性値を持つ。同一のタグ名を持つ開始タグおよび終了タグに囲まれた領域のことを項目と呼ぶ。ある項目の中に別の項目を含むことが可能である（すなわち階層構造が可能である）。

【0085】例えば、XML文書と呼ばれる構造化文書では、タグの開始を表す符号は

<

であり、タグの終了を表す符号は

>

であり、開始タグは

<タグ名 属性名=属性値…>

終了タグは

</タグ名>

空タグは

<タグ名 属性名=属性値…/>

でそれぞれ定義される。

【0086】また、構造化文書には、文書型定義（Document Type Definition; DTD）と呼ばれる、文書構造を定義する文書を関連づけることが出来るものもある。DTDは、ある構造化文書の中で、どのようなタグを用いることが出来るか、それらのタグの階層関係はどのようになっているかなどを定義し、構造化文書の論理構造を明確に定義するものである。

【0087】以下、第1の実施形態では、本発明の文書内重要項目抽出装置について、重要項目抽出装置の具体的な構成と動作例、重要項目の抽出方法の設定手順例、抽出した重要項目の表示方法例、抽出する重要項目の変更方法例について、それぞれ詳細に説明する。

【0088】第2の実施形態では、付加項目生成機能を備えた文書内重要項目抽出装置を構造化文書を送受信するデータ送受信装置で実現した例、検索機能を備えた文書内重要項目抽出装置を構造化文書を送受信するデータ送受信装置で実現した例、付加項目生成機能と検索機能を備えた文書内重要項目抽出装置を構造化文書を送受信するデータ送受信装置で実現した例、付加項目生成機能を備えた文書内重要項目抽出装置を構造化文書を送受信するデータ中継装置で実現した例、検索機能を備えた文書内重要項目抽出装置を構造化文書を送受信するデータ中継装置で実現した例について、それぞれ詳細に説明する。

【0089】第3の実施形態では、抽出した重要項目をユーザ毎に保存する例、ユーザ毎に抽出した重要項目を複数のユーザ間で共有する例、ユーザ毎に抽出した重要項目を共有する際に、アクセス許可情報を付加して重要項目に対するアクセス権を変更することを可能にする例について、それぞれ詳細に説明する。

【0090】（第1の実施形態）まず、本発明の第1の実施形態について説明する。

【0091】本実施形態における文書内重要項目抽出装置では、「重要項目」を1つあるいは複数指定する機能を設ける。各重要項目は任意の名前を付けて識別し、1つの重要項目について1つまたは複数の抽出方法を指定できる機能を設ける。「重要項目の抽出方法」の指定は、構造化文書の「タグ名」あるいは「単語」などを指定することにより行う。

【0092】本実施形態における文書内重要項目抽出装置では、指定されたタグ名や単語などが構造化文書内に含まれている場合には、構造化文書全体もしくは構造化文書中の指定されたタグの部分のみを保存する。

【0093】構造化文書を解析する場合、すでに文書構造がタグによって構造化されているので、タグを用いて容易に解析することができる。また、抽出された重要項目の範囲を特定する場合においても、重要項目を含むタグ全体を範囲とすることによって、容易に特定することができる。このため、抽出処理の負担は非常に軽量であると言える。

【0094】このようにすれば、多くの構造化文書の中から重要項目のみあるいは重要項目を含む構造化文書のみを効果的に保存することができる。

【0095】本実施形態におけるデータ送受信装置（クライアント装置）では、構造化文書を表示する際に、文書内重要項目抽出装置が抽出し保存する部分を明示的に表示する機能を設ける。

【0096】このようにすれば、ユーザは構造化文書中の重要項目を容易に識別することができる。また、逆に重要項目であるにもかかわらず表示されない場合には、文書内重要項目抽出装置が抽出に失敗していることも確認することができる。

【0097】重要項目を明示的に表示する方法としては、例えば、重要項目を他の文書中の文字色と異なる色で表示したり、他の文書中の背景色と異なる色で表示したりするなど、様々な方法がある。

【0098】本実施形態におけるデータ送受信装置では、文書内重要項目抽出装置が自動的に抽出し保存する部分を明示的に指定、追加、変更、削除できる機能を設ける。

【0099】例えば、マウスのような位置指示装置（ポインティングデバイス）等を用いて構造化文書中の領域を指定することで、その部分が重要項目であることを文書内重要項目抽出装置に指定する。さらに、他の重要項

目と同じ意味の項目があった場合には、その項目を重要項目として追加する。あるいは、重要項目として指定されている項目のうち、後に重要項目で無いと判断された項目があった場合には、重要項目としてその項目を抽出しないように変更する。また、データ送受信装置が重要項目として明示的に表示している部分が重要項目でない場合には、ユーザは該当部分の自動抽出機能を削除することを指定する。

【0100】このようにすれば、ユーザは重要項目の指定、追加、変更、削除を容易に指定することができる。

【0101】重要項目の指定をする場合、例えば、構造化文書から『電話』のタグを「電話番号」という重要項目で指定することができる。さらに「TEL」という名前のついたタグも同様に電話番号を表現する重要項目であった場合には、「電話番号」という重要項目に「TEL」のタグを追加する。また、TELのタグは電話番号を表現しないとユーザが判断した場合には、「電話番号」という重要項目のうち「TEL」のタグは抽出しないように変更する。さらに、電話番号という重要項目が不要になった場合には、「電話番号」という重要項目を削除する。

【0102】本実施形態の文書内重要項目抽出装置では、重要項目を抽出する設定を削除した場合においても、すでに抽出し保存した重要項目は保存し続けられる。ある時点で不要になった重要項目であっても、保存済みの項目は継続して保存されるので、検索などの機能を継続して使用することができる。

【0103】以下では、本実施形態をより詳しく説明する。

【0104】(実施形態1-1)図1に、本実施形態に係る文書内重要項目抽出装置を用いたネットワーク・システムの構成例を示す。

【0105】図1は、(i) WWWページ作成者101、WWWサーバ104、プロキシサーバ107およびWWWページブラウザ113で構成されるWWWシステム、(ii) 電子メール送信者102、送信側電子メールサーバ105、受信側電子メールサーバ108および電子メールブラウザ114で構成される電子メールシステム、(iii) 電子ニュース記事送信者103、送信側電子ニュースサーバ106、受信側電子ニュースサーバ109および電子ニュースブラウザ115で構成される電子ニュースシステムの3種類のシステム、および、(iv) 各システム上で交換される構造化文書より重要項目を自動的に抽出する文書内重要項目抽出装置110からなるシステム構成の例である。

【0106】また、図1の例では、ユーザは、WWWページブラウザ113、電子メールブラウザ114および電子ニュースブラウザ115が動作するクライアント112を用いて所望の操作を行う。

【0107】まず、図1のWWWシステムに関する動作

を、図2のフローチャートを参照しながら説明する。

【0108】通常、WWWサーバ104上のWWWページは、WWWページ作成者101によって作成され、WWWサーバ104に保存されている。

【0109】始めに、ユーザは、クライアント112上のアプリケーションであるWWWページブラウザ113に対して、WWWサーバ104内に保存されているWWWページを表示するように指示する(ステップS201)。このとき、直接WWWページの保存されているWWWサーバ104を指示することや、すでにWWWページブラウザ113上で表示されているWWWページに含まれるハイパーリンクを指定することができる。

【0110】ユーザからの指示を受けたWWWページブラウザ113は、プロキシサーバ107にユーザの指定するWWWページの送信要求を通知する(図1中、r117)(ステップS202)。プロキシサーバ107では、同様に、WWWサーバ104にWWWページの送信要求を通知する(r118)(ステップS203)。WWWサーバ104は、要求されたWWWページをプロキシサーバ107に送信する(r119)(ステップS204)。

【0111】プロキシサーバ107は、受信したWWWページを文書内重要項目抽出装置110に入力する(r120)(ステップS205)。文書内重要項目抽出装置110は、構造化文書であるWWWページの構造を解析し、その中にユーザの指定した重要項目が含まれているか検索する(ステップS206)。重要項目が見つかったか判断し(ステップS207)、重要項目が見つかった場合には、重要項目を抽出し記録装置111に保存するとともに(ステップS208)、重要項目が抽出されたことを示すタグを、WWWページに追加する(ステップS209)。このとき、WWWページに重要項目が含まれていない場合には、記録装置111には何も保存しない。記録装置111には例えば固定ディスク装置が用いられる。

【0112】重要項目の抽出が終了した後、文書内重要項目抽出装置110は、プロキシサーバ107にWWWページを出力する(r121)。プロキシサーバ107は、WWWページをWWWページブラウザ113に転送する(r122)(ステップS210)。

【0113】WWWページを受信したWWWページブラウザ113は、WWWページを表示する(ステップS211)。このとき、WWWページに重要項目が含まれていた場合には、抽出されたことを示すタグが文書内重要項目抽出装置110によって付加されているので、その抽出領域を例えば強調表示や反転表示や網かけ表示などによって明示的に表示する。

【0114】次に、電子メールシステムにおける動作を、図3のフローチャートを参照しながら説明する。

【0115】電子メールは、電子メール送信者102に

よって、送信側電子メールサーバ105に送信され（r123）、送信側電子メールサーバ105は、電子メールの受信者であるユーザの受信側電子メールサーバ108に電子メールを送信する（r124）。

【0116】ユーザは、クライアント112上の電子メールブラウザ114に対して、受信側電子メールサーバ108に保存されている電子メールを表示するように指示する（ステップS301）。

【0117】ユーザからの指示を受けた電子メールブラウザ114は、受信側メールサーバ108に電子メールの送信要求を通知する（r125）（ステップS302）。

【0118】受信側電子メールサーバ108は、送信要求のあった電子メールを文書内重要項目抽出装置110に入力する（r126）（ステップS303）。文書内重要項目抽出装置110は、構造化文書である電子メールの構造を解析し、その中にユーザの指定した重要項目が含まれているか検索する（ステップS304）。重要項目が見つかったか判断し（ステップS305）、重要項目が見つかった場合には、重要項目を抽出し記録装置111に保存するとともに（ステップS306）、重要項目が抽出されたことを示すタグを、電子メールに追加する（ステップS307）。このとき、電子メールに重要項目が含まれていない場合には、記録装置111には何も保存しない。

【0119】重要項目の抽出が終了した後、文書内重要項目抽出装置110は、受信側電子メールサーバ108に電子メールを出力する（r127）。受信側電子メールサーバ108は、電子メールを電子メールブラウザ114に転送する（r128）（ステップS308）。

【0120】電子メールを受信した電子メールブラウザ114は、電子メールを表示する（ステップS309）。このとき、電子メールに重要項目が含まれていた場合には、抽出されたことを示すタグが文書内重要項目抽出装置110によって付加されているので、その抽出領域を例えば強調表示や反転表示や網かけ表示などによって明示的に表示する。

【0121】次に、電子ニュースシステムにおける動作を、図4のフローチャートを参照しながら説明する。

【0122】電子ニュース記事は、電子ニュース記事送信者103によって、送信側電子ニュースサーバ106に投稿される（r129）。投稿されたニュース記事は、送信側電子ニュースサーバ106からユーザの受信側電子ニュースサーバ109に転送される（r130）。

【0123】ユーザは、クライアント112上の電子ニュースブラウザ115に対して、受信側電子ニュースサーバ109内に保存されている電子ニュース記事を表示するように指示する（ステップS401）。

【0124】ユーザからの指示を受けた電子ニュースブ

ブラウザ115は、電子ニュースサーバ109に電子ニュース記事の送信要求を通知する（r131）（ステップS402）。

【0125】受信側電子ニュースサーバ109は、送信要求のあった電子ニュース記事を文書内重要項目抽出装置110に入力する（r132）（ステップS403）。文書内重要項目抽出装置110は、構造化文書である電子ニュース記事の構造を解析し、その中にユーザの指定した重要項目が含まれているか検索する（ステップS404）。重要項目が見つかったか判断し（ステップS405）、重要項目が見つかった場合には、重要項目を抽出し記録装置111に保存するとともに（ステップS406）、重要項目が抽出されたことを示すタグを、電子ニュース記事に追加する（ステップS407）。このとき、電子メール記事に重要項目が含まれていない場合には、記録装置111には何も保存しない。

【0126】重要項目の抽出が終了した後、文書内重要項目抽出装置110は、受信側電子ニュースサーバ109に電子ニュース記事を出力する（r133）。電子ニュースサーバ109は、電子ニュース記事を電子ニュースブラウザ115に転送する（r134）（ステップS408）。

【0127】電子ニュース記事を受信した電子ニュースブラウザ115は、電子ニュース記事を表示する（ステップS409）。このとき、電子ニュース記事に重要項目が含まれていた場合には、抽出されたことを示すタグが文書内重要項目抽出装置110によって付加されているので、その抽出領域を例えば強調表示や反転表示や網かけ表示などによって明示的に表示する。

【0128】以上のようにして、構造化文書は、サーバからユーザの使用するクライアントへ送信され、その途中で文書内重要項目抽出装置が動作し、所望の重要項目を自動的に抽出し、保存する。

【0129】なお、図1では、文書内重要項目抽出装置110をプロキシサーバ107と受信側電子メールサーバ108と受信側電子ニュースサーバ109で共用しているが、文書内重要項目抽出装置110をそれぞれに専用に設けるようにしてもよい。

【0130】また、例えばプロキシサーバ107と受信側電子メールサーバ108と受信側電子ニュースサーバ109を用いないシステムにおいて、文書内重要項目抽出装置110をクライアント112と接続するようにしてもよい。また、クライアント112が複数存在する場合に、それらで文書内重要項目抽出装置110を共用するようにしてもよいし、クライアント112ごとに専用に文書内重要項目抽出装置110を設けるようにしてもよい。

【0131】また、図1では、文書内重要項目抽出装置110は独立した装置として示されているが、文書内重要項目抽出装置110が、各データ中継サーバ（図1で

は、プロキシサーバ107と受信側電子メールサーバ108と受信側電子ニュースサーバ109)のそれぞれに実装するようにしてもよい。この場合、図2～図4の手順において、文書内重要項目抽出装置110とデータ中継サーバとの間のやり取りが、例えば文書内重要項目抽出プロセスとデータ中継サーバプロセスとの間の通信になる。

【0132】また、文書内重要項目抽出装置110をデータ送受信装置(図1ではクライアント112)に実装するようにしてもよい。この場合、図2～図4の手順において、文書内重要項目抽出装置110とデータ中継サーバとの間のやり取りの代わりに、例えば文書内重要項目抽出プロセスとブラウザプロセスとの間での通信が行われる。また、この場合、図1のようにデータ中継装置(図1では、プロキシサーバ107と受信側電子メールサーバ108と受信側電子ニュースサーバ109)を利用することも、データ中継装置を利用しないことも可能である。

【0133】なお、上記のバリエーションについては、後述する第2、第3の実施形態に対しても同様のことが成立する。

【0134】次に、図5および図6を参照しながら「重要項目の抽出方法」を設定するための構成と方法について説明する。

【0135】まず、図5に、重要項目の抽出方法を設定するための、文書内重要項目抽出装置110の抽出機能部分およびデータ送受信装置(クライアント装置)に関する構成の一例を示す。

【0136】文書内重要項目抽出装置110は、類語データベース501、抽出方法データベース502、重要項目データベース503を持っている。

【0137】類語データベース501には、重要項目として設定される一般的な用語の類語が格納されている。例えば、「会議」の類語として「打合せ」や「ミーティング」という単語が格納されている。

【0138】抽出方法データベース502には、ユーザが設定した「重要項目の抽出方法」が格納され、文書内重要項目抽出装置110が構造化文書より重要項目を抽出する際に参照される。

【0139】重要項目データベース503には、文書内重要項目抽出装置110により自動的に抽出された重要項目が、構造化された形式で、かつ、後から検索できるように、格納されている。

【0140】ユーザは、データ送受信装置であるクライアント112内の重要項目抽出方法設定アプリケーション506を用いて、文書内重要項目抽出装置110と通信し、重要項目の抽出方法を設定する。重要項目の抽出には、構造化文書の構造を決定するタグの名称を用いる方法と、構造化文書の全文から対応する単語を検索する方法との二通りを行うことができる。これらの方法の一

方を用いて抽出することも、双方を同時に用いて抽出することもできる。

【0141】重要項目の抽出方法は、重要項目の「分類」とその「抽出方法」とによって構成される。複数の重要項目の分類を指定することができるとともに、1つの重要項目の分類に対して複数の抽出方法を指定することができる。

【0142】次に、図6を参照しながら「重要項目の抽出方法」を設定するための手順について説明する。

【0143】ユーザは、重要項目抽出方法設定アプリケーション506を用いて重要項目の抽出方法を設定する。図6の設定手順では、はじめに、重要項目の「分類」を指定する(ステップS601)。分類には、例えば「会議通知」という名称をつけることができる。次に、1つまたは複数の抽出方法を設定する。

【0144】ステップS602～S608は所望の抽出方法を設定するためのループ処理であり、設定したい抽出方法の種類数だけループする。

【0145】まず、抽出方法を指定する(ステップS602)。

【0146】構造化文書の構造を決定するタグの名称を用いた方法で抽出を行う場合には、構造化文書内の抽出タグ名を指定する(ステップS603)。このとき、抽出タグ名の類語が類語データベース501に存在する場合(ステップS605)には、全ての類語を抽出タグ名として抽出する方法を指定することもできる(ステップS606)。類語を用いることが指定された場合には、文書内重要項目抽出装置110が構造化文書内の重要項目を抽出する際に、抽出タグ名として類語全てを用いて検索する(ステップS607)。

【0147】構造化文書の全文から対応する単語を検索する方法で抽出を行う場合には、抽出単語を指定する(ステップS604)。このとき、抽出単語の類語が類語データベース501に存在する場合(ステップS605)には、全ての類語を検索する抽出単語として指定することもできる(ステップS606)。類語を用いることが指定された場合には、文書内重要項目抽出装置110が構造化文書内の重要項目を抽出する際に、抽出単語として類語全てを用いて検索する(ステップS607)。

【0148】また、構造化文書の抽出タグ名を指定する際に、抽出する構造化文書の参照するDTDが指定されている場合には、DTDによって構造化文書内に含まれるタグ名の種類が決まるので、構造化文書内に含まれる可能性のあるタグのみを抽出タグ名に指定できるようにするのが好ましい。このようにすれば、重要項目のタグ名の設定の誤りを防ぐことができる。

【0149】なお、他の抽出方法を繰り返し指定することができる(ステップS608)。

【0150】さて、上記のようにして重要項目の分類お

よび該分類に対する1または複数の抽出方法を設定すると、重要項目抽出方法設定アプリケーション506は、文書内重要項目抽出装置110に、その設定を通知する(ステップS609)。文書内重要項目抽出装置110は、その設定を抽出方法データベース502に格納する(ステップS610)。

【0151】なお、ステップS601以降を繰り返すことによって、複数の分類を設定することができる。

【0152】次に、図5および図7を参照しながら文書内重要項目抽出装置の重要項目の抽出手順について説明する。

【0153】図7に、文書内重要項目抽出装置110の重要項目の抽出手順を示す。なお、この抽出手順の説明においては、図8に示す構造化文書の例、図9に示す抽出処理途中の構造化文書の状態の例、図10に示す抽出方法の設定例、図11に示す類語データベースの例、図12に示す抽出結果の例も併せて用いる。なお、図8(a)、図9(a)、図10(a)、図12(a)が対応しており、また図8(b)、図9(b)、図10(b)、図12(b)が対応している。

【0154】文書内重要項目抽出装置110の入力には、WWWページ、電子メール、電子ニュース記事がテキストファイルであるので、テキストファイルまたはテキストを格納したメモリを用いる。

【0155】入力されたテキストは、通常、図8(a)に例示する構造化文書801あるいは図8(b)に例示する構造化文書802のように、タグを用いて構造化されている。そこで、タグを用いて、図9(a)に例示する文書構造901や図9(b)に例示する文書構造902のように、文書構造を解析する(ステップS701)。

【0156】続いて、文書内重要項目抽出装置110は、ユーザにより設定された重要項目抽出方法、例えば図10(a)に例示する重要項目抽出方法1001や図10(b)に例示する重要項目抽出方法1002、を抽出方法データベース502より読み出し(ステップS702)、重要項目の抽出を行う。このとき、重要項目抽出方法が、重要項目抽出方法1002のように複数指定されていた場合には、全ての重要項目抽出方法に関して繰り返し抽出を行う(ステップS716)。

【0157】続いて、抽出方法として構造化文書のタグの名称を用いた方法が指定されているか判断する(ステップS703)。重要項目抽出方法1001のように抽出方法として構造化文書の構造を決定するタグの名称を用いた方法が指定されている場合には、抽出タグ名が解析された文書構造に含まれるか検索する(ステップS704)。本具体例においては、文書構造901には抽出方法に指定されている抽出タグ名である<meeting>が含まれているので(ステップS706)、抽出タグ<meeting>およびその下位層の構造1201

を重要項目として抽出し、重要項目データベース503に格納する(ステップS714)。

【0158】重要項目抽出方法1002のように抽出方法として構造化文書の全文より対応する単語を検索する方法が指定されている場合には、抽出単語が解析された文書構造に含まれるか検索する(ステップS705)。本具体例においては、文書構造902には抽出方法に指定されている抽出単語である「価格」が含まれていないので、重要項目データベース503には何も格納しない(ステップS707)。しかし、本具体例においては、抽出方法に類語を用いることが指定されているので(ステップS709)、類語による検索を行う(ステップS711)。図11に例示する類語データベース1101を参照すると、文書構造902には指定されている抽出単語の類語(1101)のうち「特価」という類語が含まれているので、抽出単語である特価を含むタグ<コメント>およびその下位層の構造1202を重要項目として抽出し、重要項目データベース503に格納する(ステップS715)。

【0159】以上の処理は、全ての重要項目について検索が行われるまで繰り返し実行される(ステップS717)。

【0160】次に、図5、図10、図13、図14を参照しながら文書内重要項目抽出装置の重要項目の格納手順を説明する。

【0161】文書内重要項目抽出装置110が重要項目を格納する重要項目データベース503には、構造化されたデータを格納することのできるデータベースを用いる。例えば、オブジェクト指向データベースを用いることができる。

【0162】図13に、重要項目データベース503の内部構造の一例を示す。なお、図13の内容は、図8に例示した2つの構造化文書から、図10に例示した抽出方法の設定で分類「会議」と分類「価格」に対する重要項目を抽出し、併せて分類「商品名」についてタグ名=<商品名>で重要項目を抽出した結果に対応するものである。

【0163】重要項目データベース1315には、重要項目抽出方法の設定の際に用いた重要項目の分類すなわち「会議」1302、「商品名」1308、「価格」1311を格納しておく。

【0164】抽出方法データベース502に格納された重要項目抽出方法によって文書内重要項目抽出装置110が抽出した重要項目は、抽出に用いられた重要項目抽出方法の「分類」の下に格納すると効果的である。このように格納することによって、各重要項目抽出方法によって抽出された重要項目を区別することが可能になる。例えば、図10(a)に例示した抽出方法1001で抽出された重要項目は、分類「会議」1302の下に格納する。

【0165】また、同一構造化文書より抽出された重要項目には、各々の重要項目を相互に関係付けるリンクを付加するようにするのが好ましい。例えば、分類「商品名」1308に分類された重要項目1309と分類「価格」1311に分類された重要項目1312とが同一構造化文書より抽出された場合には、重要項目1309と重要項目1312とを、リンク1314によって関係付けることができる。このように重要項目間をリンクにより関係付けることによって、ある構造化文書内から異なる重要項目抽出方法によって格納された重要項目間の関係を保存することができるので、重要項目を検索する際に、複数の重要項目が同一構造化文書内に含まれていたという関係を利用することができる。

【0166】図14に、重要項目データベース503の内部構造の他の例を示す。図14の例は、図13にさらに抽出された重要項目が追加された例を示している。

【0167】図14の例では、他の構造化文書から分類「会議」1302に関する重要項目1325が抽出され、分類「会議」1302の下に重要項目1325が追加・格納されている様子を表している。また、他の構造化文書から分類「商品名」1308に関する重要項目1320と分類「価格」1311に関する重要項目1322が抽出され、それぞれ分類「会議」1302や分類「価格」1311の下に追加・格納されている様子を表している。さらに、分類「商品名」1308に分類された重要項目1320と分類「価格」1311に分類された重要項目1322とが同一構造化文書より抽出されたので、これら重要項目1320と重要項目1322とがリンク1324によって関係付けられている様子を表している。

【0168】（実施形態1-2）次に、データ送受信装置（図1のクライアント112）における、ブラウザ（図1のWWWブラウザ113、電子メールブラウザ114および電子ニュースブラウザ115）の構造化文書の表示の方法と、重要項目の抽出方法の指定・追加・変更・削除の方法についてそれぞれ説明する。

【0169】WWWブラウザ113と電子メールブラウザ114と電子ニュースブラウザ115での各操作は同じであるので、以下では、WWWブラウザ113上での操作を例にとって説明する。

【0170】まず、図15、図16、図17を参照しながら、新たに重要項目を指定するための操作例について説明する。

【0171】図15に、WWWブラウザ113で重要項目が含まれていない構造化文書を表示する例を示す。図15では、構造化文書の題目1601、構造化文書の内容1602、構造化文書の記書き部分1603は、いずれも重要項目を含んでいないため、構造化文書が作成された状態のままで表示される。

【0172】このとき、構造化文書の記書き部分160

3をWWWブラウザ113上で新たに重要項目として指定するものとする。

【0173】ユーザは、例えば位置指示装置であるマウスポインタ1604を用いて、構造化文書の記書き部分1603を領域指定し、重要項目指定ボタン1605を押す。図16に、記書き部分1603が領域指定された様子を示す。

【0174】領域が指定され重要項目指定ボタン1605が押されると、WWWブラウザ113は、図17に例示する重要項目指定画面を表示し、構造化文書内からユーザにより領域指定された部分のタグを検索し抽出名入力部分1802を設定する。ユーザは重要項目名入力部分1801に重要項目名を入力する。同様に、抽出方法としてタグ名を用いる場合にはタグ抽出指定ボタン1803を押す。単語を用いる場合には単語抽出指定ボタン1804を押す。さらに、重要項目の保存方法として、重要項目を含む構造化文書全体を保存する場合には全文保存指定ボタン1805を押す。重要項目であるタグのみを保存する場合にはタグ保存指定ボタン1806を押す。最後に、入力内容決定ボタン1807を押す。入力内容を決定し、文書内重要項目抽出装置110に重要項目の抽出を指定する。

【0175】次に、図16、図18を参照しながら、構造化文書の記書き部分1603を既存の重要項目に追加するための操作例について説明する。

【0176】ユーザは、図15の表示状態から、図16のように例えば位置指示装置であるマウスポインタ1604を用いて構造化文書の記書き部分1603を領域指定し、重要項目追加ボタン1606を押す。

【0177】領域が指定され重要項目追加ボタン1606が押されると、WWWブラウザ113は、図18に示す重要項目追加画面を表示し、構造化文書内からユーザにより領域指定された部分のタグを検索し抽出名入力部分2002を設定する。ユーザは重要項目名入力部分2001に既存の重要項目名を入力する。同様に、抽出方法としてタグ名を用いる場合にはタグ抽出指定ボタン2003を押す。単語を用いる場合には単語抽出指定ボタン2004を押す。さらに、重要項目の保存方法として、重要項目を含む構造化文書全体を保存する場合には全文保存指定ボタン2005を押す。重要項目であるタグのみを保存する場合にはタグ保存指定ボタン2006を押す。最後に、入力内容決定ボタン2007を押す。入力内容を決定し、文書内重要項目抽出装置110に重要項目の抽出を指定する。

【0178】次に、図16、図19を参照しながら、構造化文書の記書き部分1603が既存の重要項目である場合にその重要項目を変更するための操作例について説明する。

【0179】ユーザは、図15の表示状態から、図16のように例えば位置指示装置であるマウスポインタ16



04を用いて構造化文書の記書き部分1603を領域指定し、重要項目変更ボタン1607を押す。

【0180】領域が指定され重要項目変更ボタン1607が押されると、WWWブラウザ113は、図19に示す重要項目変更画面を表示し、該当する重要項目名を重要項目名入力部分2201に設定する。同様に、抽出名を抽出名入力部分2202に設定し、抽出方法としてタグ名を用いている場合にはタグ抽出指定ボタン2203を設定し、単語を用いている場合には単語抽出指定ボタン2204を設定する。さらに、重要項目の保存方法として、重要項目を含む構造化文書全体を保存する場合には全文保存指定ボタン2205を設定し、重要項目であるタグのみを保存する場合にはタグ保存指定ボタン2206を設定する。

【0181】ユーザは、該当重要項目の設定内容を変更し、入力内容決定ボタン2207を押し、入力内容を決し、文書内重要項目抽出装置110に重要項目の抽出方法の変更を指定する。

【0182】次に、図16、図20を参照しながら、構造化文書の記書き部分1603が既存の重要項目である場合にその重要項目を削除するための操作例について説明する。

【0183】ユーザは、図15の表示状態から、図16のように例えば位置指示装置であるマウスポインタ1604を用いて、構造化文書の記書き部分1603を領域指定し、重要項目削除ボタン1608を押す。

【0184】領域が指定され重要項目削除ボタン1608が押されると、WWWブラウザ113は、図20に示す重要項目変更画面を表示し、該当する重要項目名を重要項目名入力部分2401に設定する。同様に、抽出名を抽出名入力部分2402に設定し、抽出方法としてタグ名を用いている場合にはタグ抽出指定ボタン2403を設定し、単語を用いている場合には単語抽出指定ボタン2404を設定する。さらに、重要項目の保存方法として、重要項目を含む構造化文書全体を保存する場合には全文保存指定ボタン2405を設定し、重要項目であるタグのみを保存する場合にはタグ保存指定ボタン2406を設定する。

【0185】ユーザは、該当重要項目の設定内容を確認し削除を行う場合には、入力内容決定ボタン2407を押し、文書内重要項目抽出装置110に重要項目の抽出の削除を指定する。

【0186】本実施形態によれば、本文書内重要項目抽出装置をWWWサーバなどの構造化文書配信装置から配信される構造化文書を閲覧する際に用いると、文書内重要項目抽出装置は、ユーザに代わって、受信した構造化文書内からユーザにとって重要である項目を自動的に抽出し、記憶装置に保存する。構造化文書内の重要項目の抽出、保存作業は、文書内重要項目抽出装置によって自動的に行われるので、ユーザは、受信した構造化文書

とに自ら重要項目を抽出、保存する煩わしさから開放され、抽出し忘れることも保存し忘れることもなくなる。

【0187】また、本実施形態に係る文書内重要項目抽出装置をWWWサーバなどの構造化文書配信装置から配信される構造化文書を閲覧する際に用いると、ユーザが利用するブラウザなどのデータ送受信装置において、抽出された重要項目が確認でき、ユーザの意図通りに重要項目の抽出、保存が行われているかどうかを確認することが出来る。また、データ送受信装置において、重要項目抽出方法設定アプリケーションを用いることにより、抽出すべき重要項目の指定、追加、変更、削除を行うことが出来、ユーザは必要に応じて、適切な重要項目を自動的に抽出することが可能となる。

【0188】(第2の実施形態)次に、本発明の第2の実施形態について説明する。

【0189】第1の実施形態では、文書内重要項目抽出装置の持つ基本的な機能として、構造化文書から重要項目を抽出して保存する機能について説明したが、本実施形態では、構造化文書から抽出した重要項目を保存する機能に加えて、当該抽出した重要項目に関連のある付加項目を新たに生成し、当該抽出した重要項目と当該生成した付加項目とを対応付けてもしくは組にして保存する機能についても説明する。重要項目に関連のある付加項目としては、様々なものが生成可能であるが、例えば、構造化文書から重要項目を抽出した日時を生成し保存しておくことにより、後から情報の新しさを容易に知ることができる。

【0190】また、本実施形態では、文書内重要項目抽出装置に、検索機能を設けることについても説明する。検索機能が提供されていれば、ユーザは以前に保存した重要項目や付加項目の中から必要な情報を効率よく探し出すことができるようになる。なお、ここでは、検索機能を第2の実施形態として説明するが、第1の実施形態で説明した文書内重要項目抽出装置に検索機能(重要項目を検索対象とする)を設けることもできる。

【0191】第1の実施形態でも述べたように、本文書内重要項目抽出装置は、構造化文書を送受信するデータ送受信装置(図1では、クライアント112)で実現する方式と構造化文書を送受信するデータ中継装置(図1では、プロキシサーバ107と受信側電子メールサーバ108と受信側電子ニュースサーバ109)で実現する方式が考えられる。また、後者の代わりに、データ中継装置の外部に、本文書内重要項目抽出装置を独立した装置として設け、データ中継装置と文書内重要項目抽出装置との間で情報のやり取りを行う方法や、さらにデータ中継装置の外部に独立した装置として設けられた文書内重要項目抽出装置を複数のデータ中継装置で共有する方法もある。

【0192】さて、本文書内重要項目抽出装置は、基本的な機能として、構造化文書配信装置(図1では、WW

Wサーバ104、送信側電子メールサーバ105、送信側電子ニュースサーバ106)から通信路を経由して配信される構造化文書を受信し、受信した構造化文書内から重要項目を自動的に抽出し、記録装置に保存する機能を有する。

【0193】構造化文書内の項目のうち、どのような条件を満たす項目が文書内重要項目抽出装置のユーザにとって重要であるかについては、ユーザが文書内重要項目抽出装置に対して指示をする。ユーザから指示が行われたら、以後、文書内重要項目抽出装置は、指示内容にしたがって、構造化文書内から重要項目の抽出、記録装置への保存を行う。

【0194】ユーザから文書内重要項目抽出装置に対して、構造化文書内の重要項目を満たすべき条件を指示する方法は、例えば、以下に示すような方法が考えられる。

- (1) 構造化文書内の「タグ名」で与える方法
- (2) 構造化文書内の「タグの属性名」で与える方法
- (3) 構造化文書内の「タグの属性値」で与える方法
- (4) 構造化文書内の「タグ以外のテキスト中に含まれるキーワード」で与える方法
- (5) 上記の(1)～(4)の一部または全部を組み合わせた方法

構造化文書内の項目のうち、どのような条件を満たす項目がユーザにとって重要であるかをまとめた情報は、テーブルの形式で文書内重要項目抽出装置内の記憶装置に保存し、このテーブルを参照したい場合には、適宜、記憶装置から読み出すものとする。このテーブルのことを「構造化文書内重要項目条件テーブル」と呼ぶものとする。

【0195】上記の(1)のように、構造化文書内の重要項目がタグ名で与えられた場合には、文書内重要項目抽出装置は、当該与えられたタグ名を持つ項目を構造化文書内から抽出し、記録装置に保存する。

【0196】上記の(2)のように、構造化文書内の重要項目がタグの属性名で与えられた場合には、文書内重要項目抽出装置は、当該与えられたタグの属性名を持つ項目を抽出し、記録装置に保存する。

【0197】上記の(3)のように、構造化文書内の重要項目がタグの属性値で与えられた場合には、文書内重要項目抽出装置は、当該与えられたタグの属性値を持つ項目を抽出し、記録装置に保存する。

【0198】上記の(4)のように、構造化文書内の重要項目がタグ以外のテキスト中に含まれるキーワードで与えられた場合には、文書内重要項目抽出装置は、当該与えられたキーワードを含む項目を抽出し、記録装置に保存する。

【0199】いずれの場合にも、記録装置に保存するデータは、構造化文書の項目の集合になっており、構造化文書になっている。

【0200】次に、本実施形態に係る、付加項目を生成し保存する機能を有する文書内重要項目抽出装置は、構造化文書配信装置から通信路を経由して配信される構造化文書を受信し、受信した構造化文書内から重要項目を自動的に抽出し、また抽出した重要項目に関連のある付加項目を自動的に生成し、両者を対応付けてもしくは組にして記録装置に保存する。

【0201】どのような条件を満たす項目を付加項目として生成するかについては、ユーザが文書内重要項目抽出装置に(付加項目指示機能を利用して)指示する。ユーザから指示が行われたら、以後、文書内重要項目抽出装置は、指示内容にしたがって、構造化文書内から付加項目の生成、記録装置への保存を行う。

【0202】ユーザから文書内重要項目抽出装置に対して、抽出された構造化文書内の重要項目に関連する付加項目を指示する方法は、例えば、以下に示すような方法が考えられる。

- (1) 付加項目は「重要項目を抽出した日時」であると指示する方法
- (2) 付加項目は「構造化文書が存在する位置」であると指示する方法
- (3) 付加項目は「抽出された重要項目以外の項目が存在する位置」であると指示する方法
- (4) 上記の(1)～(3)の一部または全部を組み合わせた方法

構造化文書内の重要項目に関連する付加項目をまとめた情報は、テーブルの形式で文書内重要項目抽出装置内の記録装置に保存し、このテーブルを参照したい場合には、適宜、記録装置から読み出すものとする。このテーブルのことを「構造化文書内重要項目関連付加項目条件テーブル」と呼ぶことにする。

【0203】上記の(1)のように、構造化文書内の重要項目に関連のある付加項目が、構造化文書の重要項目を抽出した日時である場合には、その日時を、記録装置に保存する項目形式のデータすなわちタグ付きのデータ形式で生成し、記録装置に保存する。

【0204】上記の(2)のように、構造化文書内の重要項目に関連のある付加項目が、構造化文書が存在する位置である場合には、対象とする構造化文書がどの構造化文書配信装置内のどの構造化文書であるかを検索し、記憶装置に保存する項目形式のデータすなわちタグ付きのデータを生成し、記録装置に保存する。

【0205】上記の(3)のように、構造化文書内の重要項目に関連のある付加項目が、抽出した重要項目以外の項目が存在する位置である場合には、対象とする構造化文書内の項目がどの構造化文書配信装置内のどの構造化文書内のどの項目であるかを検索し、記録装置に保存する項目形式のデータすなわちタグ付きのデータを生成し、記録装置に保存する。

【0206】いずれの場合にも、記録装置に保存するデ



ータは、タグ付きデータ（項目）の集合になっており、構造化文書になっている。

【0207】次に、本実施形態に係る、検索機能を有する文書内重要項目抽出装置は、記録装置に保存してある重要項目の検索（付加項目を生成し保存する機能を有しない場合）または重要項目や付加項目の検索（付加項目を生成し保存する機能を有する場合）を行うことが可能である。

【0208】ユーザが、検索したい重要項目が満たすべき条件を、文書内重要項目抽出装置に指示すると、文書内重要項目抽出装置は、検索条件に合致する重要項目を、記録装置にすでに保存されている重要項目の中から検索する。合致する重要項目が見つかった場合には当該重要項目を表示し、見つからなかった場合には見つからなかった旨を表示する。

【0209】文書内重要項目抽出装置の記録装置に保存されている項目は、以下に示す2種類がある。

- (1) 受信した構造化文書内から抽出した重要項目
- (2) 抽出された重要項目に関連する付加項目

検索の対象となる項目は、上記の2種類の項目の両方である。

【0210】ユーザから文書内重要項目抽出装置に検索する項目を指示する方法は、例えば以下に示す方法が考えられる。

- (1) 項目の「タグ名」で与える方法
- (2) 項目の「タグの属性名」で与える方法
- (3) 項目の「タグの属性値」で与える方法
- (4) 項目の「タグ以外のテキスト中に含まれるキーワード」で与える方法
- (5) 上記の(1)～(4)の一部または全部を組み合わせた方法

上記の(1)のように、ユーザの検索したい項目が項目のタグ名で与えられた場合には、記録装置の中に保存されている項目の中から、該与えられたタグ名を持つ項目を検索し、読み出す。

【0211】上記の(2)のように、ユーザの検索したい項目が項目のタグの属性名で与えられた場合には、記録装置の中に保存されている項目の中から、該与えられたタグの属性名を持つ項目を検索し、読み出す。

【0212】上記の(3)のように、ユーザが検索したい項目が項目のタグの属性値で与えられた場合には、記録装置の中に保存されている項目の中から、該与えられたタグの属性値を持つ項目を検索し、読み出す。

【0213】上記の(4)のように、ユーザが検索したい項目が項目のタグ以外のテキスト中に含まれるキーワードで与えられた場合には、記録装置の中に保存されている項目の中から、該与えられたキーワードを含む項目を検索し、読み出す。

【0214】次に、重要項目および付加項目を記録装置に格納する際のデータ形式について説明する。

【0215】図21に、重要項目および付加項目を格納する重要項目・付加項目データベースの一例を示す。

【0216】この例では、

- ・構造化文書の中から重要項目として、<date>タグで囲まれた会議開催日と
- ・<place>で囲まれた会議開催場所とを抽出し、抽出した重要項目に関連する付加項目として、
- ・構造化文書がもともと存在していた場所を表すURLと
- ・重要項目の抽出を行った日を作成し、それぞれ<URL>、<extractday>という名前のタグで囲んでいる。

【0217】抽出した重要項目は<main>というタグで囲み、作成した付加項目は<sub>というタグで囲んで、全体として1つの構造化文書になるようにしてある。

【0218】検索を行う際には、

- (1) 抽出した重要項目のみの検索
- (2) 作成した付加項目のみの検索
- (3) 抽出した重要項目、作成した付加情報の違いを意識しない検索

の3つを適宜使い分けることが可能である。

【0219】構造化文書がホームページの場合には、<URL>タグの中身は文字通りURLになり、構造化文書がメールやニュースである場合には、<URL>タグの中身は、メールやニュースの送信者のメールアドレスにしておく。

【0220】以下では、本実施形態の文書内重要項目抽出装置の構成・動作についてより詳しく説明する。まず、文書内重要項目抽出装置をデータ送受信装置（図1では、クライアント112）で実施する形態について説明し、次に、文書内重要項目抽出装置をデータ中継装置（図1では、プロキシサーバ107と受信側電子メールサーバ108と受信側電子ニュースサーバ109）で実施する形態について説明する。

【0221】（実施形態2-1）図22に、本実施形態（2-1）に係る、重要項目を抽出し保存する機能と付加項目を生成し保存する機能と検索機能とを有する文書内重要項目抽出装置の構成例を示す。本構成例は、文書内重要項目抽出装置をデータ送受信装置で実現したものである。

【0222】なお、重要項目を抽出し保存する機能を有する文書内重要項目抽出装置は、図22の構成から、構造化文書内重要項目関連付加項目指示制御部126、付加項目指示部150、検索指示制御部127、検索指示部160を省いた構成に相当する。

【0223】また、重要項目を抽出し保存する機能と付加項目を生成し保存する機能を有する文書内重要項目抽出装置は、図22の構成から、検索指示制御部127、検索指示部160を省いた構成に相当する。

【0224】また、重要項目を抽出し保存する機能と検索機能を有する文書内重要項目抽出装置は、図22の構成から、構造化文書内重要項目関連付加項目指示制御部126、付加項目指示部150を省いた構成に相当する。

【0225】なお、これら一部の機能を有しないものについては、その詳細な説明は省略する。

【0226】図22に示されるように、本文書内重要項目抽出装置110は、構造化文書処理部120、構造化文書閲覧部130、重要項目指示部140、付加項目指示部150、検索指示部160を備え、通信路170を介してネットワークと接続されている。

【0227】構造化文書処理部120は、構造化文書要求解析部121、構造化文書受信部122、構造化文書内重要項目抽出部123、構造化文書内重要項目関連付加項目生成部124、構造化文書内重要項目指示制御部125、構造化文書内重要項目関連付加項目指示制御部126、検索指示制御部127、記憶部制御部128、記憶部129を備えている。

【0228】構造化文書閲覧部130は、構造化文書要求入力部131、構造化文書表示部132を備えている。

【0229】重要項目指示部140は、重要項目指示入力部141、重要項目指示実行結果表示部142を備えている。

【0230】付加項目指示部150は、付加項目指示入力部151、付加項目指示実行結果表示部152を備えている。

【0231】検索指示部160は、検索指示入力部161、検索結果表示部162を備えている。

【0232】＜構造化文書閲覧部130＞構造化文書閲覧部130は、ユーザからの構造化文書閲覧要求を受け付け、構造化文書処理部120に対して構造化文書配信要求を送信する。また、構造化文書処理部120から送信される構造化文書を受信、表示する。

【0233】構造化文書要求入力部131は、ユーザが構造化文書閲覧要求を入力する部分である。ユーザが構造化文書閲覧要求を行う方法には、URLを指定する方法などがある。

【0234】構造化文書表示部132は、構造化文書処理部120から構造化文書閲覧部130に送信される構造化文書を表示する部分である。

【0235】＜重要項目指示部140＞重要項目指示部140は、構造化文書処理部120が受信する構造化文書内の項目のうち、どのような条件を満たす項目が文書内重要項目抽出装置のユーザにとって重要であるかを、ユーザから構造化文書処理部120に指示するための部分である。

【0236】ユーザは、構造化文書処理部120が受信する構造化文書内の項目のうち、どのような条件を満た

す項目が自分にとって重要であるかを、重要項目指示入力部141に入力する。

【0237】ユーザが重要項目指示入力部141に入力する内容には、例えば以下のようなものが考えられる。

(1) 構造化文書内重要項目条件テーブルに新たな条件を追加する。

(2) 構造化文書内重要項目条件テーブルにすでに保存されている条件を、新しい条件に変更する。

(3) 構造化文書内重要項目条件テーブルにすでに保存されている条件のうち、あるものを削除する。

(4) 構造化文書内重要項目条件テーブルの内容を、重要項目指示実行結果表示部142に表示する。

【0238】重要項目指示実行結果表示部142は、構造化文書処理部120から重要項目指示部140に送信されたデータを表示する。構造化文書処理部120から重要項目指示部140に送信されるデータとは、具体的には以下の2つが考えられる。

(1) ユーザが重要項目指示入力部141に入力した指示が成功したかどうかを示す情報

(2) 現在構造化文書処理部120に登録されている構造化文書内重要項目条件テーブルの内容

＜付加項目指示部150＞付加項目指示部150は、構造化文書処理部120が構造化文書内から抽出した重要項目に関連する付加項目がどのような条件を満たすべきかを、ユーザから構造化文書処理部120に指示するための部分である。

【0239】ユーザは、構造化文書処理部120が構造化文書内から抽出した重要項目に関連する付加項目がどのような条件を満たすべきかを、付加項目指示入力部151に入力する。

【0240】ユーザが付加項目指示入力部151に入力する内容は、例えば以下のようなものが考えられる。

(1) 構造化文書内重要項目関連付加項目条件テーブルに、新たな条件を追加する。

(2) 構造化文書内重要項目関連付加項目条件テーブルにすでに保存されている条件を、新しい条件に変更する。

(3) 構造化文書内重要項目関連付加項目条件テーブルにすでに保存されている条件のうち、あるものを削除する。

(4) 構造化文書内重要項目関連付加項目条件テーブルの内容を、付加項目指示実行結果表示部152に表示する。

【0241】付加項目指示実行結果表示部152は、構造化文書処理部120から付加項目指示部150に送信されたデータを表示する。構造化文書処理部120から付加項目指示部150に送信されるデータとは、具体的には以下の2つが考えられる。

(1) ユーザが付加項目指示入力部151に入力した指示が成功したかどうかを示す情報

(2) 現在構造化文書処理部120に登録されている構

造化文書内重要項目関連付加項目条件テーブルの内容  
 <検索指示部160>検索指示部160は、ユーザが検索したいと思っている項目が満たすべき条件を、ユーザから構造化文書処理部120に指示するための部分である。

【0242】ユーザは、構造化文書処理部120内に保存されている項目のうち、自分が検索したいと思っている項目が満たすべき条件を検索指示入力部161に入力する。

【0243】検索結果表示部162は、構造化文書処理部120から検索指示部160に送信されたデータを表示する。

【0244】構造化文書処理部120から検索指示部160に送信されるデータとは、具体的には以下の2つが考えられる。

(1) ユーザが検索指示入力部161に入力した指示が成功したかどうかを示す情報

(2) 構造化文書処理部120で検索を行い、条件に合致した項目の組

<通信路170>通信路170は、構造化文書閲覧部130から構造化文書処理部120に対して送信された構造化文書配信要求を中継し、WWWサーバや電子メールサーバや電子ニュースサーバなどの構造化文書配信装置に対して送信する。また、WWWサーバや電子メールサーバや電子ニュースサーバなどの構造化文書配信装置から送信された構造化文書の中継し、構造化文書処理部120に対して送信する。

【0245】なお、図1のようにデータ中継装置(図1では、プロキシサーバ107と受信側電子メールサーバ108と受信側電子ニュースサーバ109)を利用することも、データ中継装置を利用しないことも可能である。

【0246】<構造化文書処理部120>構造化文書処理部120は、WWWサーバや電子メールサーバなどの構造化文書配信装置から送信された構造化文書を受信し、受信した構造化文書を構造化文書閲覧部130に転送するとともに、受信した構造化文書内から重要項目を抽出し、抽出した重要項目に関連する付加項目を生成し、両者を組にした上で記憶部129に保存する。同じ文書名を持つ構造化文書を以前にも受信している場合には、再度抽出、生成を再度行う必要があるかどうかの確認を行う必要がある。また、構造化文書処理部120は、ユーザが重要項目指示部140や付加項目指示部150を利用して指定する方法にしたがって、重要項目と付加項目の追加、変更、削除を行う。さらに、構造化文書処理部120は、ユーザが検索したい項目を検索指示部160を利用して構造化文書処理部120に通知すると、ユーザの要求に合致する項目を記憶部129に保存されている項目の中から検索し、読み出した上で、検索指示部160に転送する。

【0247】<<構造化文書要求解析部121>>構造化文書要求解析部121は、構造化文書閲覧部130から送信された構造化文書配信要求を受信し、要求された構造化文書がどの構造化文書配信装置上に存在するかを解析し、通信路160を経由して、要求された構造化文書が存在するWWWサーバや電子メールサーバなどの構造化文書配信装置に対して、または要求された構造化文書が存在する構造化文書配信装置と構造化文書内重要項目抽出装置110との間に位置するプロキシサーバなどのデータ中継装置に対して、構造化文書配信要求を送信する。

【0248】<<構造化文書受信部122>>構造化文書受信部122は、WWWサーバや電子メールサーバなどの構造化文書配信装置から通信路110を経由して構造化文書処理部120に送信された構造化文書を受信し、受信した構造化文書の複写を行う。複写した構造化文書の一方は、構造化文書閲覧部130に転送し、他方は構造化文書内重要項目抽出部123に転送する。

【0249】<<構造化文書内重要項目抽出部123>>構造化文書内重要項目抽出部123は、構造化文書受信部122から送信された構造化文書を受信し、構造化文書内重要項目条件テーブルの内容にしたがい、構造化文書内から重要項目の抽出を行う。構造化文書内重要項目抽出部123は、構造化文書内重要項目条件テーブルの内容を一時保持するための記憶領域を持つ。

【0250】ここでもし、構造化文書内重要項目抽出部123が構造化文書内重要項目条件テーブルの内容を保持している場合には、このテーブルが最新のものであるかどうかを記憶部制御部128に問い合わせ、もし最新のものでない場合には、記憶部制御部128に対して最新のテーブル送信を要求し、構造化文書内重要項目抽出部123内に最新の構造化文書内重要項目条件テーブルを保持する。また、構造化文書内重要項目抽出部123が構造化文書内重要項目条件テーブルの内容を保持していない場合には、記憶部制御部128に対して最新のテーブル送信を要求し、構造化文書内重要項目抽出部123内に最新の構造化文書内重要項目条件テーブルを保持する。

【0251】受信した構造化文書が以前に受信した構造化文書と同じ文書名を持つ構造化文書であった場合には、再度重要項目の抽出を行う必要があるかどうかを確認する。確認とは、具体的には以下に示す3つの方法が考えられる。

(1) 今回受信した構造化文書の内容が1番最近更新された日時が、前回構造化文書から重要項目の抽出を行った日時よりも後であるかどうかを確認する。

(2) 前回重要項目の抽出を行った日時よりも後に、構造化文書内重要項目条件テーブルの内容に変化があったかどうかを確認する。

(3) 上記の(1)と(2)の両方を行う。

【0252】確認を行った結果、再度重要項目の抽出を行う必要があると判断された場合には、構造化文書内重要項目抽出部123は、構造化文書内重要項目条件テーブルの内容にしたがって、構造化文書からユーザにとって重要な項目をすべて抽出する。

【0253】抽出作業が終了したら、構造化文書内重要項目抽出部123は、抽出を行う前の構造化文書と、抽出した重要項目すべてを、構造化文書内重要項目関連付加項目生成部124に送信する。

【0254】<<構造化文書内重要項目関連付加項目生成部124>>構造化文書内重要項目関連付加項目生成部124は、構造化文書内重要項目抽出部123から送信された構造化文書と構造化文書から抽出された重要項目とを受信し、構造化文書内重要項目関連付加項目条件テーブルの内容にしたがい、付加項目の生成を行う。構造化文書内重要項目関連付加項目生成部124は、構造化文書内重要項目関連付加項目条件テーブルの内容を一時保持するための記憶領域を持つ。

【0255】ここでもし、構造化文書内重要項目関連付加項目生成部124が構造化文書内重要項目関連付加項目条件テーブルの内容を保持している場合には、このテーブルが最新のものであるかどうかを記憶部制御部128に問い合わせ、もし最新のものでない場合には、記憶部制御部128に対して最新のテーブル送信を要求し、構造化文書内重要項目関連付加項目生成部124内に最新の構造化文書内重要項目関連付加項目条件テーブルの内容を保持する。また、構造化文書内重要項目関連付加項目生成部124が構造化文書内重要項目関連付加項目条件テーブルの内容を保持していない場合には、記憶部制御部128に対して最新のテーブル送信を要求し、構造化文書内重要項目関連付加項目生成部124内に最新の構造化文書内重要項目関連付加項目条件テーブルの内容を保持する。

【0256】受信した構造化文書が以前に受信した構造化文書と同じ文書名を持つ構造化文書であった場合には、再度重要項目の抽出を行う必要があるかどうかを確認する。確認とは、具体的には以下に示す3つの方法が考えられる。

(1) 今回受信した構造化文書の内容が1番最近更新された日時が、前回構造化文書から重要項目の抽出を行った日時よりも後であるかどうかを確認する。

(2) 前回重要項目の抽出を行った日時よりも後に、構造化文書内重要項目条件テーブルの内容に変化があったかどうかを確認する。

(3) 上記の(1)と(2)の両方を行う。

【0257】ただし、(1)に示す確認作業が構造化文書内重要項目抽出部123ですでになされている場合には、再度行う必要はない。

【0258】確認を行った結果、再度付加項目の生成を行う必要があると判断された場合には、構造化文書内重

要項目関連付加項目生成部124は、構造化文書内重要項目関連付加項目条件テーブルの内容にしたがって、構造化文書内重要項目抽出部123で抽出した重要項目に関連する付加項目を生成する。

【0259】生成作業が終了したら、構造化文書内重要項目関連付加項目生成部124は、構造化文書内重要項目抽出部123が抽出した構造化文書内重要項目と、構造化文書内重要項目関連付加項目生成部124が生成した構造化文書内重要項目関連付加項目とをまとめて、記憶部制御部128に送信し、記憶部制御部128に対して、抽出、生成した項目の記憶部129への保存を指示する。

【0260】<<構造化文書内重要項目指示制御部125>>構造化文書内重要項目指示制御部125は、重要項目指示部140から送信される指示内容を受信し、受信した内容にしたがって構造化文書内重要項目条件テーブルの操作を行う。

【0261】構造化文書内重要項目指示制御部125で、重要項目指示部140から送信された指示内容を受信し、受信した内容の解析を行う。重要項目指示部140から送信される指示内容とは、例えば以下のようなものが考えられる。

(1) 構造化文書内重要項目条件テーブルに新たな条件を追加する。

(2) 構造化文書内重要項目条件テーブルにすでに保存されている条件を、新しい条件に変更する。

(3) 構造化文書内重要項目条件テーブルにすでに保存されている条件のうち、あるものを削除する。

(4) 構造化文書内重要項目条件テーブルの内容を、重要項目指示実行結果出力部141に表示する。

【0262】重要項目指示部140から送信される指示内容が、構造化文書内重要項目条件テーブルへの新たな条件の追加であった場合には、構造化文書内重要項目条件テーブルへ条件の追加を記憶部制御部128に対して指示する。構造化文書内重要項目指示制御部125から記憶部制御部128に対して行った指示に対する返答を受信し、重要項目指示部140に送信する。

【0263】構造化文書内重要項目指示制御部125から記憶部制御部128に対して行った指示に対する返答とは具体的には、例えば、以下に示す2つが考えられる。

(1) 新しい条件の追加に成功した。

(2) 記憶部129に空き領域がないなどの理由で、新しい条件の追加に失敗した。

【0264】重要項目指示部140から送信される指示内容が、構造化文書内重要項目条件テーブルにすでに保存されている条件の変更であった場合には、構造化文書内重要項目条件テーブルへ条件の変更を記憶部制御部128に対して指示する。構造化文書内重要項目指示制御部125から記憶部制御部128に対して行った指示に

に対する返答を受信し、重要項目指示部 140 に送信する。

【0265】構造化文書内重要項目指示制御部 125 から記憶部制御部 128 に対して行った指示に対する返答とは具体的には、例えば、以下に示す 2 つが考えられる。

- (1) 条件の変更に成功した。
- (2) テーブル上に変更しようとしている条件が見つからないなどの理由で、新しい条件の追加に失敗した。

【0266】重要項目指示部 140 から送信される指示内容が、構造化文書内重要項目条件テーブルにすでに保存されている条件の削除であった場合には、構造化文書内重要項目条件テーブルから条件を削除することを記憶部制御部 128 に対して指示する。構造化文書内重要項目指示制御部 125 から記憶部制御部 128 に対して行った指示に対する返答を受信し、重要項目指示部 140 に送信する。

【0267】構造化文書内重要項目指示制御部 125 から記憶部制御部 128 に対して行った指示に対する返答とは、具体的には、例えば、以下に示す 2 つが考えられる。

- (1) 条件の削除に成功した。
- (2) 削除すべき条件がテーブル上に見つからないなどの理由で、条件の変更に失敗した。

【0268】重要項目指示部 140 から送信される指示内容が、構造化文書内重要項目条件テーブルの内容の重要項目指示実行結果出力部 141 への表示であった場合には、構造化文書内重要項目条件テーブルの記憶部 129 からの読み出しと、重要項目指示部 140 への送信を記憶部制御部 128 に対して指示する。構造化文書内重要項目指示制御部 125 から記憶部制御部 128 に対して行った指示に対する返答を受信し、重要項目指示部 140 に送信する。

【0269】構造化文書内重要項目指示制御部 125 から記憶部制御部 128 に対して行った指示に対する返答をと、具体的には、例えば、以下に示す 2 つが考えられる。

- (1) テーブルの内容
- (2) 記憶部 129 からのテーブルの読み出しに失敗したなどの理由で、テーブルの読み出し、転送に失敗した旨を示す情報

<<構造化文書内重要項目関連付加項目指示制御部 126>>構造化文書内重要項目関連付加項目指示制御部 126 は、付加項目指示部 150 から送信される指示内容を受信し、受信した内容にしたがって構造化文書内重要項目関連付加項目条件テーブルの操作を行う。

【0270】構造化文書内重要項目関連付加項目指示制御部 126 で、付加項目指示部 150 から送信された指示内容を受信し、受信した内容の解析を行う。

【0271】付加項目指示部 150 から送信される指示

内容とは、例えば以下のようなものが考えられる。

(1) 構造化文書内重要項目関連付加項目条件テーブルに新たな条件を追加する。

(2) 構造化文書内重要項目関連付加項目条件テーブルにすでに保存されている条件を、新しい条件に変更する。

(3) 構造化文書内重要項目関連付加項目条件テーブルにすでに保存されている条件のうち、あるものを削除する。

(4) 構造化文書内重要項目関連付加項目条件テーブルの内容を、付加項目指示実行結果表示部 152 に表示する。

【0272】付加項目指示部 150 から送信される指示内容が、構造化文書内重要項目関連付加項目条件テーブルへの新たな条件の追加であった場合には、構造化文書内重要項目条件テーブルへ条件の追加を記憶部制御部 128 に対して指示する。構造化文書内重要項目関連付加項目指示制御部 126 から記憶部制御部 128 に対して行った指示に対する返答を受信し、付加項目指示部 150 に送信する。

【0273】構造化文書内重要項目関連付加項目指示制御部 126 から記憶部制御部 128 に対して行った指示に対する返答とは具体的には、例えば、以下に示す 2 つが考えられる。

- (1) 新しい条件の追加に成功した。
- (2) 記憶部 129 に空き領域がないなどの理由で、新しい条件の追加に失敗した。

【0274】付加項目指示部 150 から送信される指示内容が、構造化文書内重要項目関連付加項目条件テーブルにすでに保存されている条件の変更であった場合には、構造化文書内重要項目条件テーブルの条件の変更を記憶部制御部 128 に対して指示する。構造化文書内重要項目関連付加項目指示制御部 126 から記憶部制御部 128 に対して行った指示に対する返答を受信し、付加項目指示部 150 に送信する。

【0275】構造化文書内重要項目関連付加項目指示制御部 126 から記憶部制御部 128 に対して行った指示に対する返答とは具体的には、例えば、以下に示す 2 つが考えられる。

- (1) 新しい条件の追加に成功した。
- (2) テーブル上に変更しようとしている条件が見つからないなどの理由で、新しい条件の追加に失敗した。

【0276】付加項目指示部 150 から送信される指示内容が、構造化文書内重要項目関連付加項目条件テーブルにすでに保存されている条件の削除であった場合には、構造化文書内重要項目条件テーブルからの条件の削除を記憶部制御部 128 に対して指示する。構造化文書内重要項目関連付加項目指示制御部 126 から記憶部制御部 128 に対して行った指示に対する返答を受信し、付加項目指示部 150 に送信する。

【0277】構造化文書内重要項目関連付加項目指示制御部126から記憶部制御部128に対して行った指示に対する返答とは具体的には、例えば、以下に示す2つが考えられる。

- (1) 新しい条件の追加に成功した。
- (2) テーブル上に削除しようとしている条件が見つからないなどの理由で、新しい条件の追加に失敗した。

【0278】付加項目指示部150から送信される指示内容が、構造化文書内重要項目関連付加項目条件テーブルの内容の付加項目指示実行結果表示部152への表示であった場合には、構造化文書内重要項目条件テーブルの記憶部129からの読み出しと、付加項目指示部150への送信を記憶部制御部128に対して指示する。構造化文書内重要項目関連付加項目指示制御部126から記憶部制御部128に対して行った指示に対する返答を受信し、付加項目指示部150に送信する。

【0279】構造化文書内重要項目関連付加項目指示制御部126から記憶部制御部128に対して行った指示に対する返答とは、具体的には、例えば、以下に示す2つが考えられる。

- (1) テーブルの内容
- (2) 記憶部129からのテーブルの読み出しに失敗したなどの理由で、テーブルの読み出し、転送に失敗した旨を示す情報

<<検索指示制御部127>>検索指示制御部127は、検索指示部160から送信される検索条件を受信し、検索条件の解析をする。検索条件に合致する項目の記憶部129からの読み出しを記憶部制御部128に指示し、指示した結果記憶部制御部128から送信される情報を検索指示部160に送信する。

【0280】検索指示部160から送信される検索条件は、例えば、以下に示すような方法で与えられる。

- (1) 重要項目のタグ名で与えられる。
- (2) 重要項目のタグの属性名で与えられる。
- (3) 重要項目のタグの属性値で与えられる。
- (4) 重要項目のタグ以外のテキスト中に含まれるキーワードで与えられる。
- (5) 上記の(1)～(4)の一部または全部の組み合わせで与える。

【0281】また、指示した結果記憶部制御部128から送信される情報とは、具体的には以下の2つが考えられる。

- (1) 検索条件に合致する項目の組
- (2) 検索条件に合致する項目が存在しないことを示す情報

<<記憶部制御部128>>記憶部制御部128は、構造化文書内重要項目関連付加項目生成部124、構造化文書内重要項目指示制御部125、構造化文書内重要項目関連付加項目指示制御部126、検索指示制御部127から送信される、記憶部129からのデータの読み出

し、転送要求や、記憶部129へのデータの書き込み要求を受け付け、実行する。

【0282】記憶部制御部128は、これから記憶部129に書き込むデータや記憶部129から読み出したデータを一時保持するための記憶領域を持つ。記憶部129からの読み出し、転送を指示されたデータがすでに記憶領域上に存在する場合には、再度記憶部129からの読み出しは行わずに、転送のみを行えばよい。

【0283】記憶部129からの読み出し、記憶部129への書き込みに失敗した場合には、その旨を通知する。

【0284】<<記憶部129>>記憶部129は、データの保存を行う。データとは具体的には以下のものである。

- (1) 構造化文書から抽出した重要項目
- (2) 構造化文書から抽出した重要項目に関連する付加項目

- (3) 構造化文書内重要項目条件テーブル
  - (4) 構造化文書内重要項目関連付加項目条件テーブル
- さて、次に、ユーザにとって重要な項目が満たすべき条件を新たに構造化文書内重要項目条件テーブルに追加したい場合に行う手続きについて説明する。

【0285】図23に、この手続きのフローチャートの一例を示す。

【0286】ユーザは、重要項目指示部140の重要項目指示入力部141に、自分にとって重要な構造化文書内の項目が満たすべき条件と、この条件を新たに構造化文書内重要項目条件テーブルに追加したい旨を入力する。入力された指示内容は、重要項目指示部140から構造化文書処理部120に転送される(ステップS101)。

【0287】構造化文書処理部120の構造化文書内重要項目指示制御部125は、重要項目指示部140から送信された指示内容を受信する。構造化文書内重要項目指示制御部125が受信した指示内容を解析し、指示内容が、構造化文書内重要項目条件テーブルへの新しい条件の追加であることを知る(ステップS102)。

【0288】重要項目指示部140は、引き続き、記憶部制御部128に対して、構造化文書内重要項目条件テーブルへの新しい条件の追加を指示する(ステップS103)。

【0289】記憶部制御部128は、構造化文書内重要項目条件テーブルへの新しい条件の追加を行う。追加に成功した場合には成功した旨を、追加に失敗した場合には失敗した旨を、構造化文書内重要項目指示制御部125に送信する(ステップS104、S105、S106)。

【0290】構造化文書内重要項目指示制御部125は、記憶部制御部128から受信した内容を重要項目指示部140に送信する(ステップS107)。



【0291】重要項目指示部140は、構造化文書内重要項目指示制御部125から送信された内容を受信し、重要項目指示実行結果表示部142に表示する（ステップS108）。

【0292】なお、構造化文書内重要項目条件テーブルの条件の変更、削除、構造化文書内重要項目条件テーブル表示に関しては省略する。

【0293】次に、構造化文書から抽出した重要項目に関連する付加項目が満たすべき条件を新たに構造化文書内重要項目条件テーブルに追加したい場合に行う手続きについて説明する。

【0294】図24に、この手続きのフローチャートの一例を示す。

【0295】ユーザは、付加項目指示部150の付加項目指示入力部151に、構造化文書から抽出した重要項目に関連する付加項目が満たすべき条件と、この条件を新たに構造化文書内重要項目関連付加項目条件テーブルに追加したい旨を入力する。入力された指示内容は、付加項目指示部150から構造化文書処理部120に転送される（ステップS201）。

【0296】構造化文書処理部120の構造化文書内重要項目関連付加項目指示制御部126は、付加項目指示部150から送信された指示内容を受信する。構造化文書内重要項目関連付加項目指示制御部126が受信した指示内容を解析し、指示内容が構造化文書内重要項目関連付加項目条件テーブルへの新しい条件の追加であることを知る（ステップS202）。

【0297】付加項目指示部150は、引き続き、記憶部制御部128に対して、構造化文書内重要項目関連付加項目条件テーブルへの新しい条件の追加を指示する（ステップS203）。

【0298】記憶部制御部128は、構造化文書内重要項目関連付加項目条件テーブルへの新しい条件の追加を行う。追加に成功した場合には成功した旨を、追加に失敗した場合には失敗した旨を、構造化文書内重要項目関連付加項目指示制御部126に送信する（ステップS204、S205、S206）。

【0299】構造化文書内重要項目関連付加項目指示制御部126は、記憶部制御部128から受信した内容を付加項目指示部150に送信する（ステップS207）。

【0300】付加項目指示部150は、構造化文書内重要項目関連付加項目指示制御部126から送信された内容を受信し、付加項目指示実行結果表示部152に表示する（ステップS208）。

【0301】なお、構造化文書内重要項目関連付加項目条件テーブルの条件の変更、削除、構造化文書内重要項目関連付加項目条件テーブル表示に関しては省略する。

【0302】次に、ユーザが構造化文書の閲覧要求を出してから構造化文書内重要項目抽出装置に重要項目と付

加項目が保存されるまでに行う手続きについて説明する。

【0303】図25に、この手続きのフローチャートの一例を示す。

【0304】構造化文書閲覧部130のユーザは、自分が閲覧したいと思っている構造化文書の閲覧要求を構造化文書要求入力部131に入力する。構造化文書閲覧部130は、ユーザの入力を受け付け、構造化文書配信要求を構造化文書処理部120に対して送信する（ステップS301）。

【0305】構造化文書処理部120の構造化文書要求解析部121は、構造化文書閲覧部130から送信された構造化文書配信要求を受信し、要求された構造化文書がどの構造化文書配信装置上に存在するかを解析し、通信路160を経由して、要求された構造化文書が存在するWWWサーバや電子メールサーバなどの構造化文書配信装置に対して、または要求された構造化文書が存在するWWWサーバや電子メールサーバなどの構造化文書配信装置と構造化文書処理部120の間に位置する構造化文書送受信装置に対して、構造化文書配信要求を送信する（ステップS302）。

【0306】構造化文書配信要求を受信した構造化文書配信装置は、要求された構造化文書を通信路110などを経由して構造化文書処理部120に送信する（ステップS303）。

【0307】構造化文書処理部120内の構造化文書受信部122は、WWWサーバや電子メールサーバなどの構造化文書配信装置から通信路170を経由して構造化文書処理部120に送信された構造化文書を受信し、受信した構造化文書の複写を行う。複写した構造化文書の1方は、構造化文書閲覧部130に転送し、他方は構造化文書内重要項目抽出部123に送信する（ステップS304）。

【0308】構造化文書閲覧部130は、構造化文書処理部120から送信された構造化文書を受信し、構造化文書表示部132に表示する。また、構造化文書内重要項目抽出部123は、構造化文書受信部122から送信された構造化文書を受信する（ステップS305）。

【0309】次に、最新の構造化文書内重要項目条件テーブルを構造化文書内重要項目抽出部123内に用意する。受信した構造化文書から重要項目の抽出を行う必要があるかどうかの確認を行う（ステップS307）。

【0310】もし重要項目の抽出を行う必要がある場合には、最新の構造化文書内重要項目条件テーブルの内容にしたがって構造化文書内から重要項目をすべて抽出する。抽出作業が終了したら、抽出を行う前の構造化文書と、抽出した重要項目すべてを、構造化文書内重要項目関連付加項目生成部124に送信する（ステップS308）。

【0311】構造化文書内重要項目関連付加項目生成部

124は、構造化文書内重要項目抽出部123から送信された、構造化文書と、抽出した重要項目すべてを受信する。次に、最新の構造化文書内重要項目条件テーブルを構造化文書内重要項目関連付加項目生成部124内に用意する。付加項目の生成を行う必要があるかどうかの確認を行う(ステップS309)。

【0312】もし付加項目の生成を行う必要がある場合には、最新の構造化文書内重要項目条件テーブルの内容にしたがって、重要項目に関連のある付加項目をすべて抽出する(ステップS310)。

【0313】生成作業が終了したら、構造化文書内重要項目関連付加項目生成部124は、構造化文書内重要項目抽出部123が抽出した構造化文書内重要項目と、構造化文書内重要項目関連付加項目生成部124が生成した構造化文書内重要項目関連付加項目をまとめて、記憶部制御部128に送信し、記憶部制御部128に対して、これらの項目の記憶部129への保存を指示する(ステップS311)。

【0314】記憶部制御部128は、構造化文書内重要項目抽出部123が抽出した構造化文書内重要項目と、構造化文書内重要項目関連付加項目生成部124が生成した構造化文書内重要項目関連付加項目を、これから記憶部129に書き込むデータを一時保持するための記憶領域に受信し、受信した項目を記憶部129へ保存する。もし記憶部129へ保存しようとしている項目に対応する構造化文書名が、すでに記憶部129へ保存されている項目に対応する構造化文書名と同じであった場合、以前に保存した項目を保持したまま別の記憶領域に新しい項目を保存するか、それとも以前に項目が保存してあった領域に新しい項目を上書きするか両方の方法が考えられる(ステップS312)。

【0315】次に、ユーザが文書内重要項目抽出装置内に格納されている項目の検索を行う際に行う手続きについて説明する。

【0316】図26に、この手続きのフローチャートの一例を示す。

【0317】ユーザが、構造化文書処理部120内に保存されている項目のうち、自分が検索したいと思っている項目が満たすべき条件を検索指示部160の検索指示入力部161に入力する。検索指示部160は、ユーザが入力した内容を構造化文書処理部120に送信する(ステップS401)。

【0318】検索指示制御部127は、検索指示部160から送信される検索条件を受信し、検索条件の解析をする(ステップS402)。

【0319】検索条件に合致する項目が記憶されているかどうか判定する(ステップS403)。

【0320】検索条件に合致する項目が記憶されている場合、この項目の記憶部129からの読み出しを記憶部制御部128に指示する。検索条件に合致する項目が存

在する場合、検索指示部160から読み出しを指示された項目は、データを一時保持するための記憶領域上にあるかも知れないし、記憶部129上にしか存在しないかも知れない。記憶部制御部128は、検索指示部160から読み出しを指示された項目を、必要があれば記憶部129から読み出し、検索指示制御部127に送信する(ステップS404)。

【0321】検索条件に合致する項目が存在しない場合には、その旨を検索指示制御部127に送信する(ステップS405)。

【0322】検索指示制御部127は受信した内容を検索指示部160に送信する(ステップS406)。

【0323】検索指示部160は送信された内容を受信し、検索結果表示部162に表示する(ステップS407)。

【0324】(実施形態2-2)図27に、本実施形態(2-2)に係る、重要項目を抽出し保存する機能と付加項目を生成し保存する機能と検索機能を有する文書内重要項目抽出装置を含む、構造化文書保存・検索システムの構成例を示す。本構成例は、文書内重要項目抽出装置をデータ中継装置で実現したものである。

【0325】なお、重要項目を抽出し保存する機能を有する文書内重要項目抽出装置は、図27の構成から、構造化文書内重要項目関連付加項目指示制御部226、付加項目指示部250、検索指示制御部227、検索指示部260を省いた構成に相当する。

【0326】また、重要項目を抽出し保存する機能と付加項目を生成し保存する機能を有する文書内重要項目抽出装置は、図22の構成から、検索指示制御部227、検索指示部260を省いた構成に相当する。

【0327】また、重要項目を抽出し保存する機能と検索機能を有する文書内重要項目抽出装置は、図27の構成から構造化文書内重要項目関連付加項目指示制御部226、付加項目指示部250を省いた構成に相当する。

【0328】なお、これら一部の機能を有しないものについては、その詳細な説明は省略する。

【0329】図27に示されるように、本構造化文書保存・検索システムは、文書内重要項目抽出装置220、構造化文書閲覧装置210、通信路270、構造化文書送受信装置280からなる。

【0330】文書内重要項目抽出装置220は、構造化文書閲覧装置要求解析部221、構造化文書受信部222、構造化文書内重要項目抽出部223、構造化文書内重要項目関連付加項目生成部224、構造化文書内重要項目指示制御部225、構造化文書内重要項目関連付加項目指示制御部226、検索指示制御部227、記憶部制御部228、記憶部229を備えている。

【0331】[構造化文書閲覧装置210] 構造化文書閲覧装置210は、構造化文書閲覧部230、重要項目指示部240、付加項目指示部250、検索指示部26



0を備えている。

【0332】構造化文書閲覧部230は、構造化文書要求入力部231、構造化文書表示部232を備えている。

【0333】重要項目指示部240は、重要項目指示入力部241、重要項目指示実行結果表示部242を備えている。

【0334】付加項目指示部250は、付加項目指示入力部251、付加項目指示実行結果表示部252を備えている。

【0335】検索指示部260は、検索指示入力部261、検索結果表示部262を備えている。

【0336】＜構造化文書閲覧部230＞構造化文書閲覧部230は、ユーザからの構造化文書閲覧要求を受け付け、構造化文書送受信装置280に対して構造化文書配信要求を送信する。また、構造化文書送受信装置280から送信される構造化文書を受信、表示する。

【0337】構造化文書要求入力部231は、ユーザが構造化文書閲覧要求を入力する部分である。ユーザが構造化文書閲覧要求を行う方法には、URLを指定する方法などがある。

【0338】構造化文書表示部232は、構造化文書送受信装置280から構造化文書閲覧部230に送信される構造化文書を表示する部分である。

【0339】＜重要項目指示部240＞重要項目指示部240は、文書内重要項目抽出装置220が受信する構造化文書内の項目のうち、どのような条件を満たす項目が文書内重要項目抽出装置のユーザにとって重要であるかを、ユーザから文書内重要項目抽出装置220に指示するための部分である。

【0340】ユーザは、文書内重要項目抽出装置220が受信する構造化文書内の項目のうち、どのような条件を満たす項目が自分にとって重要であるかを、重要項目指示入力部241に入力する。

【0341】重要項目指示実行結果表示部242は、構造化文書送受信装置280から重要項目指示部240に送信されたデータを表示する。

【0342】＜付加項目指示部250＞付加項目指示部250は、文書内重要項目抽出装置220が構造化文書内から抽出した重要項目に関連する付加項目がどのような条件を満たすべきかを、ユーザから文書内重要項目抽出装置220に指示するための部分である。

【0343】ユーザは、文書内重要項目抽出装置220が構造化文書内から抽出した重要項目に関連する付加項目がどのような条件を満たすべきかを付加項目指示入力部251に入力する。

【0344】付加項目指示実行結果表示部252は、構造化文書送受信装置280から付加項目指示部250に送信されたデータを表示する。

【0345】＜検索指示部260＞検索指示部260

は、ユーザが検索したいと思っている項目が満たすべき条件を、ユーザから文書内重要項目抽出装置220に指示するための部分である。

【0346】ユーザは、文書内重要項目抽出装置220内に保存されている項目のうち、自分が検索したいと思っている項目が満たすべき条件を検索指示入力部261に入力する。

【0347】検索結果表示部262は、構造化文書送受信装置280から検索指示部260に送信されたデータを表示する。

【0348】通信路270は、構造化文書閲覧部230から文書内重要項目抽出装置220に対して送信された構造化文書配信要求を中継し、WWWサーバや電子メールサーバなどの構造化文書配信装置に対して送信する。また、WWWサーバや電子メールサーバなどの構造化文書配信装置から送信された構造化文書の中継し、文書内重要項目抽出装置220に対して送信する。

【0349】〔構造化文書受信装置280〕構造化文書受信装置280は、文書内重要項目抽出装置220と構造化文書閲覧装置210の間に位置し、文書内重要項目抽出装置220と構造化文書閲覧装置210間で送受信されるデータを中継する。

【0350】〔文書内重要項目抽出装置220〕文書内重要項目抽出装置220は、WWWサーバや電子メールサーバなどの構造化文書配信装置から送信された構造化文書を受信し、受信した構造化文書を、構造化文書送受信装置280を経由して構造化文書閲覧装置210に転送すると同時に、受信した構造化文書内から重要項目を抽出し、抽出した重要項目に関連する付加項目を生成し、両者を組にした上で記憶部に保存する。同じ文書名を持つ構造化文書を以前にも受信している場合には、再度抽出、生成を再度行う必要があるのかどうかの確認を行う必要がある。また、文書内重要項目抽出装置220は、ユーザが重要項目指示部240、付加項目指示部250を利用して指定する方法にしたがって、重要項目と付加項目の追加、変更、削除を行う。さらに文書内重要項目抽出装置220は、ユーザが検索したい項目を検索指示部260を利用して文書内重要項目抽出装置220に通知すると、ユーザの要求に合致する項目を記憶部229に保存されている項目の中から検索し、読み出した上で、構造化文書送受信装置280を経由して検索指示部260に転送する。

【0351】＜＜構造化文書閲覧装置要求解析部221＞＞構造化文書閲覧装置要求解析部221は、構造化文書閲覧部230から構造化文書送受信装置280を経由して文書内重要項目抽出装置220に送信された内容を受信し、受信した内容の解析を行った上で、しかるべき場所に転送する。

【0352】構造化文書閲覧部230から文書内重要項目抽出装置220に送信される内容とは、具体的には、

以下に示す4つが考えられる。

(1) 構造化文書閲覧部230から送信された構造化文書配信要求

(2) 重要項目指示部240から送信された重要項目指示

(3) 付加項目指示部250から送信された付加項目指示

(4) 検索指示部260から送信された検索指示

構造化文書閲覧部230から文書内重要項目抽出装置220に送信された内容が、構造化文書閲覧部230から送信された構造化文書配信要求である場合、要求された構造化文書がどの構造化文書配信装置上に存在するかを解析し、通信路260を経由して、要求された構造化文書が存在するWWWサーバや電子メールサーバなどの構造化文書配信装置に対して、または要求された構造化文書が存在する構造化文書配信装置と構造化文書内重要項目抽出装置210の間に位置する構造化文書送受信に対して、構造化文書配信要求を送信する。

【0353】構造化文書閲覧部230から文書内重要項目抽出装置220に送信された内容が、重要項目指示部240から送信された重要項目指示である場合、構造化文書内重要項目指示制御部225へ転送する。

【0354】構造化文書閲覧部230から文書内重要項目抽出装置220に送信された内容が、付加項目指示部250から送信された付加項目指示である場合、構造化文書内重要項目関連付加項目指示制御部226へ転送する。

【0355】構造化文書閲覧部230から文書内重要項目抽出装置220に送信された内容が、検索指示部260から送信された検索指示である場合、検索指示制御部227へ転送する。

【0356】<<構造化文書受信部222>>構造化文書受信部222は、WWWサーバや電子メールサーバなどの構造化文書配信装置から通信路270を経由して文書内重要項目抽出装置220に送信された構造化文書を受信し、受信した構造化文書の複写を行う。複写した構造化文書の1方は、構造化文書送受信装置280を経由して構造化文書閲覧装置210に転送し、他方は構造化文書内重要項目抽出部223に転送する。

【0357】<<構造化文書内重要項目抽出部223>>構造化文書内重要項目抽出部223は、構造化文書受信部222から送信された構造化文書を受信し、構造化文書内重要項目条件テーブルの内容にしたがい、構造化文書内から重要項目の抽出を行う。構造化文書内重要項目抽出部223は、構造化文書内重要項目条件テーブルの内容を一時保持するための記憶領域を持つ。

【0358】受信した構造化文書が以前に受信した構造化文書と同じ文書名をもつ構造化文書である場合には、再度重要項目の抽出を行う必要があるかどうかを確認する。確認を行った結果、再度重要項目の抽出を行う必要

があると判断された場合には、構造化文書内重要項目抽出部223は、最新の構造化文書内重要項目条件テーブルの内容にしたがって、構造化文書からユーザにとって重要な項目をすべて抽出する。

【0359】抽出作業が終了したら、構造化文書内重要項目抽出部223は、抽出を行う前の構造化文書と、抽出した重要項目すべてを、構造化文書内重要項目関連付加項目生成部224に送信する。

【0360】<<構造化文書内重要項目関連付加項目生成部224>>構造化文書内重要項目関連付加項目生成部224は、構造化文書内重要項目抽出部223から送信された構造化文書と構造化文書から抽出された重要項目を受信し、構造化文書内重要項目関連付加項目条件テーブルの内容にしたがい、付加項目の生成を行う。構造化文書内重要項目関連付加項目生成部224は、構造化文書内重要項目関連付加項目条件テーブルの内容を一時保持するための記憶領域を持つ。

【0361】受信した構造化文書が以前に受信した構造化文書と同じ文書名を持つ構造化文書であった場合には、再度重要項目の抽出を行う必要があるかどうかを確認する。確認を行った結果、再度付加項目の生成を行う必要があると判断された場合には、構造化文書内重要項目関連付加項目生成部224は、最新の構造化文書内重要項目関連付加項目条件テーブルの内容にしたがって、構造化文書内重要項目抽出部223で抽出した重要項目に関連する付加項目を生成する。

【0362】生成作業が終了したら、構造化文書内重要項目関連付加項目生成部224は、構造化文書内重要項目抽出部223が抽出した構造化文書内重要項目と、構造化文書内重要項目関連付加項目生成部224が生成した構造化文書内重要項目関連付加項目をまとめて、記憶部制御部228に送信し、記憶部制御部228に対して、抽出、生成した項目の記憶部229への保存を指示する。

【0363】<<構造化文書内重要項目指示制御部225>>構造化文書内重要項目指示制御部225は、構造化文書閲覧装置要求解析部221から送信される指示内容を受信し、受信した内容にしたがって構造化文書内重要項目条件テーブルの操作を行う。

【0364】<<構造化文書内重要項目関連付加項目指示制御部226>>構造化文書内重要項目関連付加項目指示制御部226は、構造化文書閲覧装置要求解析部221から送信される指示内容を受信し、受信した内容にしたがって構造化文書内重要項目関連付加項目条件テーブルの操作を行う。

【0365】<<検索指示制御部227>>検索指示制御部227は、構造化文書閲覧装置要求解析部221から送信される指示内容を受信し、検索条件の解析をする。検索条件に合致する項目の記憶部229からの読み出しを記憶部制御部228に指示し、指示した結果記憶

部制御部228から送信される情報を構造化文書送受信装置280を経由して検索指示部260に送信する。

【0366】<<記憶部制御部228>>記憶部制御部228は、構造化文書内重要項目関連付加項目生成部224、構造化文書内重要項目指示制御部225、構造化文書内重要項目関連付加項目指示制御部226、検索指示制御部227から送信される、記憶部229からのデータの読み出し、転送要求や、記憶部229へのデータの書き込み要求を受け付け、実行する。

【0367】記憶部制御部228は、これから記憶部229に書き込むデータや記憶部229から読み出したデータを一時保持するための記憶領域を持つ。記憶部229からの読み出し、転送を指示されたデータがすでに記憶領域上に存在する場合には再度記憶部229からの読み出しは行わずに、転送のみを行えばよい。

【0368】記憶部229からの読み出し、記憶部229への書き込みに失敗した場合には、その旨を通知する。

【0369】<<記憶部229>>記憶部229は、データの保存を行う。データとは具体的には以下のものである。

- (1) 構造化文書から抽出した重要項目
  - (2) 構造化文書から抽出した重要項目に関連する付加項目
  - (3) 構造化文書内重要項目条件テーブル
  - (4) 構造化文書内重要項目関連付加項目条件テーブル
- 次に、ユーザにとって重要な項目が満たすべき条件を新たに構造化文書内重要項目条件テーブルに追加したい場合に行う手続きについて説明する。

【0370】図28に、この手続きのフローチャートの一例を示す。

【0371】ユーザは、重要項目指示部240の重要項目指示入力部241に、自分にとって重要な構造化文書内の項目が満たすべき条件と、この条件を新たに構造化文書内重要項目条件テーブルに追加したい旨を入力する。入力された指示内容は、重要項目指示部240から構造化文書送受信装置280を経由して文書内重要項目抽出装置220に転送される(ステップS501)。

【0372】文書内重要項目抽出装置220の構造化文書閲覧装置要求解析部221は、重要項目指示部240から送信された指示内容を受信する。構造化文書閲覧装置要求解析部221が受信した指示内容を解析し、構造化文書内重要項目指示制御部225に転送する(ステップS502)。

【0373】構造化文書内重要項目指示制御部225は、構造化文書閲覧装置要求解析部221から転送された指示内容を受信し、指示内容を解析し、指示内容が、構造化文書内重要項目条件テーブルへの新しい条件の追加であることを知る(ステップS503)。

【0374】重要項目指示部240は、引き続き、記憶

部制御部228に対して、構造化文書内重要項目条件テーブルへの新しい条件の追加を指示する(ステップS504)。

【0375】記憶部制御部228は、構造化文書内重要項目条件テーブルへの新しい条件の追加を行う。追加に成功した場合には成功した旨を、追加に失敗した場合には失敗した旨を、構造化文書内重要項目指示制御部225に送信する(ステップS505、S506、S507)。

【0376】構造化文書内重要項目指示制御部225は、記憶部制御部228から受信した内容を、構造化文書閲覧装置要求解析部221へ転送する(ステップS508)。

【0377】構造化文書閲覧装置要求解析部221は、構造化文書内重要項目指示制御部225から受信した内容を、構造化文書送受信装置280を経由して重要項目指示部240に送信する(ステップS509)。

【0378】重要項目指示部240は、構造化文書閲覧装置要求解析部221から送信された内容を受信し、重要項目指示実行結果表示部242に表示する(ステップS510)。

【0379】なお、構造化文書内重要項目条件テーブルの条件の変更、削除、構造化文書内重要項目条件テーブル表示に関しては省略する。

【0380】次に、構造化文書から抽出した重要項目に関連する付加項目が満たすべき条件を新たに構造化文書内重要項目条件テーブルに追加したい場合に行う手続きについて説明する。

【0381】図29に、この手続きのフローチャートの一例を示す。

【0382】ユーザは、付加項目指示部250の付加項目指示入力部251に、構造化文書から抽出した重要項目に関連する付加項目が満たすべき条件と、この条件を新たに構造化文書内重要項目関連付加項目条件テーブルに追加したい旨を入力する。入力された指示内容は、付加項目指示部250から構造化文書送受信装置280を経由して文書内重要項目抽出装置220に転送される(ステップS601)。

【0383】文書内重要項目抽出装置220の構造化文書閲覧装置要求解析部221は、付加項目指示部250から送信された指示内容を受信する。構造化文書閲覧装置要求解析部221が受信した指示内容を解析し、構造化文書内重要項目関連付加項目指示制御部226に転送する(ステップS602)。

【0384】構造化文書内重要項目関連付加項目指示制御部226は、構造化文書閲覧装置要求解析部221から転送された指示内容を受信し、指示内容を解析し、指示内容が、構造化文書内重要項目関連付加項目条件テーブルへの新しい条件の追加であることを知る(ステップS603)。

【0385】付加項目指示部250は引き続き、記憶部制御部228に対して、構造化文書内重要項目関連付加項目条件テーブルへの新しい条件の追加を指示する（ステップS604）。

【0386】記憶部制御部228は、構造化文書内重要項目関連付加項目条件テーブルへの新しい条件の追加を行う。追加に成功した場合には成功した旨を、追加に失敗した場合には失敗した旨を、構造化文書内重要項目関連付加項目指示制御部226に送信する（ステップS605、S606、S607）。

【0387】構造化文書内重要項目関連付加項目指示制御部226は、記憶部制御部228から受信した内容を構造化文書閲覧装置要求解析部221へ転送する（ステップS608）。

【0388】構造化文書閲覧装置要求解析部221は、構造化文書内重要項目関連付加項目指示制御部226から受信した内容を、構造化文書送受信装置280を経由して付加項目指示部250に送信する（ステップS609）。

【0389】付加項目指示部250は、構造化文書閲覧装置要求解析部221から送信された内容を受信し、付加項目指示実行結果表示部252に表示する（ステップS610）。

【0390】なお、構造化文書内重要項目関連付加項目条件テーブルの条件の変更、削除、構造化文書内重要項目関連付加項目条件テーブル表示に関しては省略する。

【0391】次に、ユーザが構造化文書の閲覧要求を出してから構造化文書内重要項目抽出装置に重要項目と付加項目が保存されるまでに行う手続きについて説明する。

【0392】図30に、この手続きのフローチャートの一例を示す。

【0393】構造化文書閲覧部230のユーザは、自分が閲覧したいと思っている構造化文書の閲覧要求を構造化文書要求入力部231に入力する。構造化文書閲覧部230は、ユーザの入力を受け付け、構造化文書配信要求を文書内重要項目抽出装置220に対して送信する（ステップS701）。

【0394】文書内重要項目抽出装置220の構造化文書閲覧装置要求解析部221は、構造化文書閲覧部230から送信された構造化文書配信要求を受信し、要求された構造化文書がどの構造化文書配信装置上に存在するかを解析し、通信路260を経由して、要求された構造化文書が存在するWWWサーバや電子メールサーバなどの構造化文書配信装置に対して、または要求された構造化文書が存在するWWWサーバや電子メールサーバなどの構造化文書配信装置と文書内重要項目抽出装置220の間に位置する構造化文書送受信装置に対して、構造化文書配信要求を送信する（ステップS702）。

【0395】構造化文書配信要求を受信した構造化文書

配信装置は、要求された構造化文書を通信路270などを経由して文書内重要項目抽出装置220に送信する（ステップS703）。

【0396】文書内重要項目抽出装置220内の構造化文書受信部222は、サーバや電子メールサーバなどの構造化文書配信装置から通信路270を経由して文書内重要項目抽出装置220に送信された構造化文書を受信し、受信した構造化文書の複写を行う。複写した構造化文書の1方は、構造化文書送受信装置280を経由して構造化文書閲覧装置210に転送する。他方は構造化文書内重要項目抽出部223に送信する（ステップS704）。

【0397】構造化文書閲覧部230は、文書内重要項目抽出装置220から構造化文書送受信装置280を経由して送信された構造化文書を受信し、構造化文書表示部232に表示する。また、構造化文書内重要項目抽出部223は、構造化文書受信部222から送信された構造化文書を受信する（ステップS705）。

【0398】次に、最新の構造化文書内重要項目条件テーブルを構造化文書内重要項目抽出部223内に用意する。受信した構造化文書から重要項目の抽出を行う必要があるかどうかの確認を行う（ステップS707）。

【0399】もし重要項目の抽出を行う必要がある場合には、最新の構造化文書内重要項目条件テーブルの内容にしたがって構造化文書内から重要項目をすべて抽出する。抽出作業が終了したら、抽出を行う前の構造化文書と、抽出した重要項目すべてを、構造化文書内重要項目関連付加項目生成部224に送信する（ステップS708）。

【0400】構造化文書内重要項目関連付加項目生成部224は、構造化文書内重要項目抽出部223から送信された、構造化文書と、抽出した重要項目すべてを受信する。次に、最新の構造化文書内重要項目条件テーブルを構造化文書内重要項目関連付加項目生成部224内に用意する。付加項目の生成を行う必要があるかどうかの確認を行う（ステップS709）。

【0401】もし付加項目の生成を行う必要がある場合には、最新の構造化文書内重要項目条件テーブルの内容にしたがって、重要項目に関連のある付加項目をすべて抽出する。（ステップS710）。

【0402】生成作業が終了したら、構造化文書内重要項目関連付加項目生成部224は、構造化文書内重要項目抽出部223が抽出した構造化文書内重要項目と、構造化文書内重要項目関連付加項目生成部224が生成した構造化文書内重要項目関連付加項目をまとめて、記憶部制御部228に送信し、記憶部制御部228に対して、これらの項目の記憶部229への保存を指示する（ステップS711）。

【0403】記憶部制御部228は、構造化文書内重要項目抽出部223が抽出した構造化文書内重要項目と、

構造化文書内重要項目関連付加項目生成部224が生成した構造化文書内重要項目関連付加項目を、これから記憶部229に書き込むデータを一時保持するための記憶領域に受信し、受信した項目を記憶部229へ保存する。もし記憶部229へ保存しようとしている項目に対応する構造化文書名が、すでに記憶部229へ保存しようとしている項目に対応する構造化文書名が、すでに記憶部229へ保存されている項目に対応する構造化文書名と同じであった場合、以前に保存した項目を保持したまま別の記憶領域に新しい項目を保存するか、それとも以前に項目が保存してあった領域に新しい項目を上書きするか両方の方法が考えられる(ステップS712)。

【0404】次に、ユーザが文書内重要項目抽出装置内に格納されている項目の検索を行う際に行う手続きについて説明する。

【0405】図31に、この手続きのフローチャートの一例を示す。

【0406】ユーザが、文書内重要項目抽出装置220内に保存されている項目のうち、自分が検索したいと思っている項目が満たすべき条件を検索指示部260の検索指示入力部261に入力する。検索指示部260は、ユーザが入力した内容を構造化文書送受信装置280を経由して文書内重要項目抽出装置220に送信する(ステップS801)。

【0407】文書内重要項目抽出装置220の構造化文書閲覧装置要求解析部221は、検索指示部260から送信された指示内容を受信する。構造化文書閲覧装置要求解析部221が受信した指示内容を解析し、検索指示制御部227に転送する(ステップS802)。

【0408】検索指示制御部227は、検索指示部260から送信される検索条件を受信し、検索条件の解析をする(ステップS803)。

【0409】検索条件に合致する項目が記憶されているかどうか判定する(ステップS804)。

【0410】検索条件に合致する項目が記憶されている場合、この項目の記憶部229からの読み出しを記憶部制御部228に指示する。検索指示部260から読み出しを指示された項目は、データを一時保持するための記憶領域上にあるかも知れないし、記憶部229上にしか存在しないかも知れない。記憶部制御部228は、検索指示部260から読み出しを指示された項目を、必要があれば記憶部229から読み出し、検索指示制御部227に送信する(ステップS805)。

【0411】検索条件に合致する項目が存在しない場合には、その旨を検索指示制御部227に送信する(ステップS806)。

【0412】検索指示制御部227は、記憶部制御部228から受信した内容を構造化文書閲覧装置要求解析部221へ転送する(ステップS807)。

【0413】構造化文書閲覧装置要求解析部221は、

検索指示制御部227から受信した内容を、構造化文書送受信装置280を経由して付加項目指示部250に送信する(ステップS808)。

【0414】検索指示部260は送信された内容を受信し、検索結果表示部262に表示する(ステップS809)。

【0415】本実施形態によれば、本文書内重要項目抽出装置をWWWサーバなどの構造化文書配信装置から配信される構造化文書を閲覧する際に用いると、文書内重要項目抽出装置は、ユーザに代わって、受信した構造化文書内からユーザにとって重要である項目を自動的に抽出し、抽出した項目に関連がある付加項目(構造化文書の重要項目を抽出した日時、構造化文書が存在する位置、構造化文書の重要項目以外の項目が存在する位置など)を生成し、両者を組にした上で記憶装置に保存する。抽出、生成、保存の作業は文書内重要項目抽出装置によって自動的になされるので、ユーザは、閲覧した構造化文書ごとに自力で重要項目を抽出、保存する煩わしさから開放され、抽出し忘れることも保存し忘れることもなくなる。

【0416】また、本実施形態に係る文書内重要項目抽出装置は、上記方法で構造化文書の重要項目を検索すると、あらかじめ保存してある構造化文書内の重要項目や構造化文書内の重要項目に関連のある付加項目の中から所望の情報を得ることが可能となる。ユーザにとって重要な項目を文書内重要項目抽出装置に指示する際、指示の仕方が不十分であったために、ユーザにとって重要である項目が構造化文書内から抽出されなかった場合、構造化文書内の重要項目に関連する付加項目として、構造化文書が存在する位置を保持していれば、再度構造化文書全体の転送を構造化文書配信装置に依頼することで、所望の情報を得ることが可能となる。

【0417】(第3の実施形態)次に、本発明の第3の実施形態について説明する。

【0418】本実施形態では、文書内重要項目抽出装置で抽出した重要項目を保存する際に、その重要項目を抽出したユーザ毎に分離して保存する。このようにすることで、後にユーザが設定した重要項目を読み出したり、検索したりするときに、重要項目を取り出しやすくなる。

【0419】このような文書内重要項目抽出装置については、実施形態3-1、3-2で説明する。なお、実施形態3-1は、文書内重要項目抽出装置をデータ送受信装置において実現した形態(ユーザの利用するクライアントで動作している構成)であり、実施形態3-2は、文書内重要項目抽出装置をデータ中継装置において実現した形態(ユーザの利用するクライアントから独立している構成)である。

【0420】また、本実施形態では、上記のように文書内重要項目抽出装置で抽出した重要項目をユーザ毎に格



納するとともに、その重要項目に対するアクセス許可リストを付加する機能を提供する。これにより、あるユーザに関して抽出され、保存されている重要項目を、異なるユーザ間で共有することが可能になる。

【0421】このような文書内重要項目抽出装置については、実施形態3-3、3-4で説明する。実施形態3-3は、文書内重要項目抽出装置をデータ送受信装置において実現した形態であり、実施形態3-4は、文書内重要項目抽出装置をデータ中継装置において実現した形態である。

【0422】また、本実施形態では、上記のようにユーザ毎に保存されている重要項目に対してアクセス許可リストを付加する機能を提供するとともに、そのアクセス許可リストを変更する機能を提供する。これにより、あるユーザに関して抽出され、保存されている重要項目を異なるユーザ間で共有している状態を変更し、適切に共有することが可能となる。

【0423】このような文書内重要項目抽出装置については、実施形態3-5、3-6で説明する。実施形態3-5は、文書内重要項目抽出装置をデータ送受信装置において実現した形態であり、実施形態3-6は、文書内重要項目抽出装置をデータ中継装置において実現した形態である。

【0424】また、本実施形態では、複数のユーザの要求によって同一の構造化文書から同一の重要項目が抽出される場合に、このユーザ間で共通の重要項目についてはこれを共有できる形態で保存するようにする機能を提供する。

【0425】このような文書内重要項目抽出装置については、実施形態3-7で説明する。実施形態3-7は、文書内重要項目抽出装置をデータ中継装置において実現した形態である。

【0426】(実施形態3-1)図32に、本実施形態(3-1)に係る文書内重要項目抽出装置の構成例を示す。

【0427】本実施形態の文書内重要項目抽出装置10400は、ユーザインタフェース部10401、ユーザ認証部10402、構造化文書要求受信部10403、構造化文書送信部10404、重要項目抽出部10405、構造化文書送受信部10406、重要項目格納部10407を備えている。

【0428】ユーザインタフェース部10401は、ユーザとのインタフェースを司る部分であり、ユーザとの情報の入出力を可能とするため、キーボード、マウスなどの入力装置、ディスプレイ、スピーカーなどの出力装置を備えている。

【0429】ユーザ認証部10402は、文書内重要項目抽出装置10400に、アクセスしてくるユーザを認証する部分である。ユーザを認証する手段としては、例えば、パスワードを用いる認証方式などが利用できる。

認証するための情報は、ユーザインタフェース部10401を介して渡される。

【0430】構造化文書要求受信部10403は、ユーザからの要求を、ユーザインタフェース部10401経由で、構造化文書送受信部10406にユーザ情報とともに渡す部分である。

【0431】構造化文書送受信部10406は、構造化文書要求受信部10403から渡されるユーザの要求に従って、必要となる構造化文書を外部に対して要求し、該当する構造化文書を受信する部分である。

【0432】重要項目抽出部10405は、構造化文書送受信部10406から、受信した構造化文書と、該構造化文書を要求したユーザ情報を受けとり、構造化文書内の重要項目を抽出し、重要項目格納部10407に格納するとともに、構造化文書を構造化文書送信部10404へ送る部分である。また、構造化文書要求受信部10403より、重要項目の要求があった場合には、重要項目格納部10407より要求された該重要項目を構造化文書送信部10404へ送る。

【0433】重要項目の抽出は、予めユーザが指定したキーワードや構造化文書に附属しているタグを基に行う方法が考えられる。例えば、構造化文書の中に、ユーザが指定したタグが含まれている部分、およびキーワードが含まれている部分が存在した場合には、その該当する部分を抽出する。あるいは、これらのタグ、キーワードが含まれている構造化文書全体を対象としても良い。抽出した情報は、後に検索が可能のように、重要項目だけでなく、構造化文書の作成された日時、構造化文書を受信した日時、どこから構造化文書を受信したかを示す場所に関する情報なども併せて記憶しても良い。

【0434】重要項目格納部10407は、重要項目抽出部10405の指示に従い、各ユーザ毎に抽出した重要項目とそれに関連する情報を格納し、読み出す部分である。重要項目格納部10407には、それを記録するための媒体が附属している。媒体としては、磁気ディスク、光ディスク、光磁気ディスク、相変化型光ディスク、CD-R、DVD-RAM、磁気テープなどが考えられる。

【0435】構造化文書送信部10404は、重要項目抽出部10405より送られてきた構造化文書を、ユーザインタフェース部10401を介してユーザへ送信する部分である。

【0436】次に、文書内重要項目抽出装置10400の動作について説明する。

【0437】図33に、文書内重要項目抽出装置10400の動作アルゴリズムの一例を示す。

【0438】ユーザインタフェース部10401がユーザからのリクエストを受け付ける(ステップS10901)。ユーザインタフェース部10401は、受け付けたリクエストを構造化文書要求受信部10403に渡す

(ステップS10902)。

【0439】構造化文書要求受信部10403は、ユーザ認証部10402を起動し、リクエストを出したユーザをユーザインタフェース部10401を起動して認証する(ステップS10903)。構造化文書要求受信部10403は、構造化文書送受信部10406に対し、ユーザ情報とユーザのリクエストを伝える(ステップS10904)。

【0440】構造化文書送受信部10406は、外部に対し、ユーザが要求した構造化文書を要求する(ステップS10905)。外部から要求した構造化文書が送信されてきたら、構造化文書送受信部10406は、受信した構造化文書とユーザ情報を重要項目抽出部10405に渡す(ステップS10906)。

【0441】重要項目抽出部10405は、受信した構造化文書に重要項目が含まれているかどうかを検索する(ステップS10907)。重要項目が含まれている場合には、重要項目を抽出し(ステップS10909)、重要項目格納部10407を起動し、そのユーザに対応した場所に抽出した重要項目を格納する(ステップS10910)。重要項目が含まれていなかった場合には、何もしない。重要項目抽出部10405は、構造化文書送信部10404に、受信した構造化文書を渡す(ステップS10911)。

【0442】構造化文書送信部10404は、ユーザインタフェース部10401に受信した構造化文書を渡し(ステップS10912)、ユーザインタフェース部10401は、受信した構造化文書をユーザに渡す(ステップS10913)。

【0443】文書内重要項目抽出装置10400は、例えば電子計算機と同じように、CPUとメモリ装置と通信装置から構成し、文書内重要項目抽出装置10400を構成する各部分10401～10407の動作を記述したプログラムをCPUで実行することにより、その機能を得ることが出来る。

【0444】また、文書内重要項目抽出装置10400を構成する各部分10401～10407は、それぞれ、もしくはその一部を、文書内重要項目抽出装置10400と同様に、CPUとメモリ装置と通信装置から構成し、それぞれの動作を記述したプログラムをCPUで実行することにより、それぞれの機能が得られ、その集合体として構成することも可能である。

【0445】以上のように、本実施形態によれば、予め、ユーザが重要と考えるキーワードや構造化文書のタグを指定しておくことにより、受信した構造化文書から、文書内重要項目抽出装置が、ユーザに代わって重要項目を抽出することが可能となり、ユーザの手間を大きく削減することが可能となる。

【0446】(実施形態3-2)図34に、本実施形態(3-2)に係る文書内重要項目抽出装置の構成例を示

す。

【0447】本実施形態の文書内重要項目抽出装置10200は、ユーザ認証部10202、構造化文書要求受信部10203、構造化文書送信部10204、重要項目抽出部10205、構造化文書送受信部10206、重要項目格納部10207を備えている。

【0448】ユーザ認証部10202は、重要項目抽出装置10200に、ブラウザなどを用いてアクセスしてくるユーザ10201を認証する部分である。ユーザを認証する手段としては、例えば、パスワードを用いる認証方式などが利用できる。

【0449】構造化文書要求受信部10203は、ユーザ10201からの要求を、構造化文書送受信部10206にユーザ情報とともに渡す部分である。

【0450】構造化文書送受信部10206は、構造化文書要求受信部10203から渡されるユーザ10201の要求に従って、必要となる構造化文書を外部に対して要求し、該当する構造化文書を受信する部分である。

【0451】重要項目抽出部10205は、構造化文書送受信部10206から、受信した構造化文書と、該構造化文書を要求したユーザ情報を受けとり、構造化文書内の重要項目を抽出し、重要項目格納部10207に格納するとともに、構造化文書を構造化文書送信部10204へ送る部分である。また、構造化文書要求受信部10203より、重要項目の要求があった場合には、重要項目格納部10207より要求された該重要項目を構造化文書送信部10204へ送る。

【0452】重要項目の抽出は、予めユーザ10201が指定したキーワードや構造化文書に附属しているタグを基に行う方法が考えられる。例えば、構造化文書の中に、ユーザ10201が指定したタグが含まれている部分、およびキーワードが含まれている部分が存在した場合には、その該当する部分を抽出する。あるいは、これらのタグ、キーワードが含まれている構造化文書全体を対象としても良い。抽出した情報は、後に検索が可能なように、重要項目だけでなく、構造化文書の作成された日時、構造化文書を受信した日時、どこから構造化文書を受信したかを示す場所に関する情報なども合わせて記憶しても良い。

【0453】重要項目格納部10207は、重要項目抽出部10205の指示に従い、各ユーザ毎に抽出した重要項目とそれに関連する情報を格納し、読み出す部分である。重要項目格納部10207には、それを記録するための媒体が附属している。媒体としては、磁気ディスク、光ディスク、光磁気ディスク、相変化型光ディスク、CD-R、DVD-RAM、磁気テープなどが考えられる。

【0454】構造化文書送信部10204は、重要項目抽出部10205より送られてきた構造化文書を、ユーザ10201へ送信する部分である。

【0455】次に、文書内重要項目抽出装置10200の動作について説明する。

【0456】図35に、文書内重要項目抽出装置10200の動作アルゴリズムの一例を示す。

【0457】構造化文書要求受信部10203は、ユーザからのリクエストを受信する(ステップS10801)。リクエストを受信した構造化文書要求受信部10203は、ユーザ認証部10202を起動し、リクエストを出したユーザを認証する(ステップS10802)。

【0458】構造化文書要求受信部10203は、認証されたユーザ情報とユーザのリクエストを構造化文書送受信部10206に伝える(ステップS10803)。

【0459】構造化文書送受信部10206は、外部に対し、ユーザが要求した構造化文書を要求する(ステップS10804)。外部から要求した構造化文書が送信されてきたら、構造化文書送受信部10206は、受信した構造化文書とユーザ情報を重要項目抽出部10205に渡す(ステップS10805)。

【0460】重要項目抽出部10205は、受信した構造化文書に重要項目が含まれているかどうかを検索する(ステップS10806)。重要項目が含まれている場合には、重要項目を抽出し(ステップS10808)、重要項目格納部10207を起動し、そのユーザに対応した場所に抽出した重要項目を格納する(ステップS10809)。重要項目が含まれていなかった場合には、何もしない。重要項目抽出部10205は、構造化文書送信部10204に、受信した構造化文書を渡す(ステップS10810)。構造化文書送信部10204は、重要項目抽出部より渡された構造化文書をユーザに送信する(ステップS10811)。

【0461】文書内重要項目抽出装置10200は、例えば電子計算機と同じように、CPUとメモリ装置と通信装置から構成し、文書内重要項目抽出装置10200を構成する各部分10201～10207の動作を記述したプログラムをCPUで実行することにより、その機能を得ることが出来る。

【0462】また、文書内重要項目抽出装置10200を構成する各部分10201～10207は、それぞれ、もしくはその一部を、文書内重要項目抽出装置10200と同様に、CPUとメモリ装置と通信装置から構成し、それぞれの動作を記述したプログラムをCPUで実行することにより、それぞれの機能が得られ、その集合体として構成することも可能である。

【0463】以上のように、本実施形態によれば、予め、文書内重要項目抽出装置を利用するユーザ毎に、ユーザが重要と考えるキーワードや構造化文書のタグを指定しておくことにより、文書内重要項目抽出装置が、送られてきた構造化文書から、ユーザ毎にユーザに代わって重要項目を抽出することが可能となり、ユーザの手間

を大きく削減することが可能となる。

【0464】(実施形態3-3)図36に、本実施形態(3-3)に係る文書内重要項目抽出装置の構成例を示す。

【0465】本実施形態の文書内重要項目抽出装置10500は、ユーザインタフェース部10501、ユーザ認証部10502、構造化文書要求受信部10503、構造化文書送信部10504、重要項目抽出部10505、構造化文書送受信部10506、重要項目格納部10507、アクセス許可判定部10508を備えている。

【0466】ユーザインタフェース部10501は、ユーザとのインタフェースを司る部分であり、ユーザとの情報の入出力を可能とするため、キーボード、マウスなどの入力装置、ディスプレイ、スピーカーなどの出力装置を備えている。

【0467】ユーザ認証部10502は、文書内重要項目抽出装置10500にアクセスしてくるユーザを認証する部分である。ユーザを認証する手段としては、例えば、パスワードを用いる認証方式などが利用できる。

【0468】構造化文書要求受信部10503は、ユーザからの要求を、構造化文書送受信部10506にユーザ情報とともに渡す部分である。

【0469】構造化文書送受信部10506は、構造化文書要求受信部10503から渡されるユーザの要求に従って、必要となる構造化文書を外部に対して要求し、該当する構造化文書を受信する部分である。

【0470】重要項目抽出部10505は、構造化文書送受信部10506から、受信した構造化文書と、該構造化文書を要求したユーザ情報を受けとり、構造化文書内の重要項目を抽出し、重要項目格納部10507に格納するとともに、構造化文書を構造化文書送信部10504へ送る部分である。また、構造化文書要求受信部10503より、重要項目の要求があった場合には、重要項目格納部10307より要求された該重要項目を構造化文書送信部10504へ送る。

【0471】重要項目の抽出は、予めユーザが指定したキーワードや構造化文書に附属しているタグを基に行う方法が考えられる。例えば、構造化文書の中に、ユーザが指定したタグが含まれている部分、およびキーワードが含まれている部分が存在した場合には、その該当する部分を抽出する。あるいは、これらのタグ、キーワードが含まれている構造化文書全体を対象としても良い。抽出した情報は、後に検索が可能のように、重要項目だけでなく、構造化文書の作成された日時、構造化文書を受信した日時、どこから構造化文書を受信したかを示す場所に関する情報なども合わせて記憶しても良い。

【0472】重要項目格納部10507は、重要項目抽出部10505の指示に従い、各ユーザ毎に抽出した重要項目とそれに関連する情報を格納し、読み出す部分で



ある。重要項目格納部10507には、それを記録するための媒体が附属している。媒体としては、磁気ディスク、光ディスク、光磁気ディスク、相変化型光ディスク、CD-R、DVD-RAM、磁気テープなどが考えられる。

【0473】構造化文書送信部10504は、重要項目抽出部10505より送られてきた構造化文書を、ユーザへ送信する部分である。

【0474】アクセス許可判定部10508は、重要項目抽出部10505より、重要項目を読み出す際のアクセス許可を判断する部分である。アクセス許可判定部10508は、重要項目抽出部10505より、ある重要項目の読み出し要求と、読み出し要求を出したユーザ情報を与えられると、内部に持つアクセス許可リストに従い、読み出しの可、不可を判断し、判断結果を重要項目抽出部10505に返す。

【0475】ここで、図37に、抽出された重要項目と、それに対応するアクセス許可リストの格納例を示す。この例は、「会議」という重要項目を抽出したものである。「会議」に対応するキーワードとしては、「meeting」があり、文書重要項目抽出装置10500では、このキーワードが構造化文書内に含まれているかどうかを調べ、含まれている場合には抽出し、図37のような形態で格納する。このとき、重要項目だけでなく、抽出した重要項目に関するアクセス許可リストを作成し、抽出した重要項目との対応がわかるよう、格納する。この例では、「会議」に対して抽出された重要項目1は、重要項目の所有者はユーザ“foo”であり、この重要項目に対して、アクセスが許可されているのは、ユーザ“foo”とユーザ“bar”である。同様に、抽出された重要項目2に関しては、所有者はユーザ“A”であり、アクセスが許可されているのは、ユーザ“A”、“B”、“C”である。

【0476】なお、ここでは、構造化文書の要求を出したユーザが所有者となる。

【0477】次に、ユーザが外部に対して構造化文書のリクエストを出したときの文書内重要項目抽出装置10500の動作について説明する。

【0478】図38に、ユーザが外部に対して構造化文書のリクエストを出したときの文書内重要項目抽出装置10500の動作アルゴリズムの一例を示す。

【0479】ユーザインタフェース部10501がユーザからのリクエストを受け付ける（ステップS11201）。ユーザインタフェース部10501は、受け付けたリクエストを構造化文書要求受信部10503に渡す（ステップS11202）。

【0480】構造化文書要求受信部10503は、ユーザ認証部10502を起動し、リクエストを出したユーザをユーザインタフェース部10501を起動して認証する（ステップS11203）。構造化文書要求受信部

10503は、構造化文書送受信部10506に対し、ユーザ情報とユーザのリクエストを伝える（ステップS11204）。

【0481】構造化文書送受信部10506は、外部に対し、ユーザが要求した構造化文書を要求する（ステップS11205）。外部から要求した構造化文書が送信されてきたら、構造化文書送受信部10506は、受信した構造化文書とユーザ情報を重要項目抽出部10505に渡す（ステップS11206）。

【0482】重要項目抽出部10505は、受信した構造化文書に重要項目が含まれているかどうかを検索する（ステップS11207）。重要項目が含まれている場合には、重要項目を抽出し（ステップS11209）、重要項目格納部10507を起動し、そのユーザに対応した場所に抽出した重要項目を格納する（ステップS11210）。重要項目が含まれていなかった場合には、何もしない。重要項目抽出部10505は、構造化文書送信部10504に、受信した構造化文書を渡す（ステップS11211）。

【0483】構造化文書送信部10504は、ユーザインタフェース部10501に受信した構造化文書を渡し（ステップS11212）、ユーザインタフェース部10501は、受信した構造化文書をユーザに渡す（ステップS11213）。

【0484】次に、ユーザが格納されている重要項目に対してアクセス要求を出したときの文書内重要項目抽出装置10500の動作について説明する。

【0485】図39に、ユーザが格納されている重要項目に対してアクセス要求を出したときの文書内重要項目抽出装置10500の動作アルゴリズムの一例を示す。

【0486】ユーザインタフェース部10501がユーザからのリクエストを受け付ける（ステップS11301）。ユーザインタフェース部10501は、受け付けたリクエストを構造化文書要求受信部10503に渡す（ステップS11302）。

【0487】構造化文書要求受信部10503は、ユーザ認証部10502を起動し、要求を出したユーザを認証する（ステップS11303）。構造化文書要求受信部10503は、重要項目抽出部10505に対し、ユーザ情報と、アクセス要求を伝える（ステップS11304）。

【0488】重要項目抽出部10505は、アクセス許可判定部10508を起動し、ユーザが要求した重要項目へのアクセスを許可する、許可しないを判定する（ステップS11305）。アクセスが許可された場合には、重要項目抽出部10505は、重要項目格納部10507を起動し、該当する重要項目を取り出す（ステップS11307）。重要項目抽出部10505は、取り出した重要項目を構造化文書送信部10504に渡す（ステップS11309）。なお、この取り出した重要

項目も構造化文書である。アクセスが許可されなかった場合には、重要項目へのアクセスが不許可であることを示す構造化文書を構造化文書送信部10504に渡す(ステップS11308)。

【0489】構造化文書送信部10504は、渡されてきた構造化文書(取り出された重要項目またはアクセスが不許可であることを示すもの)をユーザインタフェース部10501に渡す(ステップS11310)。ユーザインタフェース部10501は、渡されてきた構造化文書をユーザに渡す(ステップS11311)。なお、上記では、アクセスが不許可であることを示す情報を構造化文書としてユーザに渡すようにしたが、アクセスが不許可であることを示す情報を構造化文書でない情報、例えば通常のメッセージ、としてユーザに渡すようにしてもよい。

【0490】なお、重要項目に対するアクセス要求を、閲覧のみか、編集も行うか、など細かく規定し、ある重要項目に対して、そのアクセス許可リストの<所有者>に登録されているユーザのみがこれを編集することを可能とし、<許可ユーザ>に登録されているユーザは閲覧のみ可能とするなどの、細かいアクセス制御を行うようにしてもよい。

【0491】また、図37の他にも、アクセス許可リストの形式には、種々のバリエーションが考えられる。

【0492】文書内重要項目抽出装置10500は、例えば電子計算機と同じように、CPUとメモリ装置と通信装置から構成し、文書内重要項目抽出装置10500を構成する各部分10501～10508の動作を記述したプログラムをCPUで実行することにより、その機能を得ることが出来る。

【0493】また、文書内重要項目抽出装置10500を構成する各部分10501～10508は、それぞれ、もしくはその一部を、文書内重要項目抽出装置10500と同様に、CPUとメモリ装置と通信装置から構成し、それぞれの動作を記述したプログラムをCPUで実行することにより、それぞれの機能が得られ、その集合体として構成することも可能である。

【0494】以上のように、本実施形態によれば、予め、文書内重要項目抽出装置を利用するユーザ毎に、ユーザが重要と考えるキーワードや構造化文書のタグを指定しておくことにより、文書内重要項目抽出装置が、送られてきた構造化文書から、ユーザ毎にユーザに代わって重要項目を抽出することが可能となり、ユーザの手間を大きく削減することが可能となる。また、抽出した重要項目に対して、アクセス制御情報を付加することにより、必要にして十分な重要項目をユーザ間で共有するのが容易になる。

【0495】(実施形態3-4)図40に、本実施形態(3-4)に係る文書内重要項目抽出装置の構成例を示す。

【0496】本実施形態の文書内重要項目抽出装置10300は、ユーザ認証部10302、構造化文書要求受信部10303、構造化文書送信部10304、重要項目抽出部10305、構造化文書送受信部10306、重要項目格納部10307、アクセス許可判定部10308を備えている。

【0497】ユーザ認証部10302は、文書内重要項目抽出装置10300に、ブラウザなどを用いてアクセスしてくるユーザ10301を認証する部分である。ユーザを認証する手段としては、例えば、パスワードを用いる認証方式などが利用できる。

【0498】構造化文書要求受信部10303は、ユーザ10301からの要求を、構造化文書送受信部10306にユーザ情報とともに渡す部分である。

【0499】構造化文書送受信部10306は、構造化文書要求受信部10303から渡されるユーザ10301の要求に従って、必要となる構造化文書を外部に対して要求し、該当する構造化文書を受信する部分である。

【0500】重要項目抽出部10305は、構造化文書送受信部10306から、受信した構造化文書と、該構造化文書を要求したユーザ情報を受けとり、構造化文書内の重要項目を抽出し、重要項目格納部10307に格納するとともに、構造化文書を構造化文書送信部10304へ送る部分である。また、構造化文書要求受信部10303より、重要項目の要求があった場合には、重要項目格納部10307より要求された該重要項目を構造化文書送信部10304へ送る。

【0501】重要項目の抽出は、予めユーザ10301が指定したキーワードや構造化文書に附属しているタグを基に行う方法が考えられる。例えば、構造化文書の中に、ユーザ10301が指定したタグが含まれている部分、およびキーワードが含まれている部分が存在した場合には、その該当する部分を抽出する。あるいは、これらのタグ、キーワードが含まれている構造化文書全体を対象としても良い。抽出した情報は、後に検索が可能なように、重要項目だけでなく、構造化文書の作成された日時、構造化文書を受信した日時、どこから構造化文書を受信したかを示す場所に関する情報なども合わせて記憶しても良い。

【0502】重要項目格納部10307は、重要項目抽出部10305の指示に従い、各ユーザ毎に抽出した重要項目とそれに関連する情報を格納し、読み出す部分である。重要項目格納部10307には、それを記録するための媒体が附属している。媒体としては、磁気ディスク、光ディスク、光磁気ディスク、相変化型光ディスク、CD-R、DVD-RAM、磁気テープなどが考えられる。

【0503】構造化文書送信部10304は、重要項目抽出部10305より送られてきた構造化文書を、ユーザ10301へ送信する部分である。

【0504】アクセス許可判定部10308は、重要項目抽出部10305より、重要項目を読み出す際のアクセス許可を判断する部分である。アクセス許可判定部10308は、重要項目抽出部10305より、ある重要項目の読み出し要求と、読み出し要求を出したユーザ情報を与えられると、内部に持つアクセス許可リストに従い、読み出しの可、不可を判断し、判断結果を重要項目抽出部10305に返す。

【0505】なお、抽出された重要項目と、それに対応するアクセス許可リストの格納例は、実施形態3-3（図37）に示す例と同様であるので、ここでの説明は省略する。

【0506】次に、ユーザが外部に対して構造化文書のリクエストを出したときの文書内重要項目抽出装置10300の動作について説明する。

【0507】図41に、ユーザが外部に対して構造化文書のリクエストを出したときの文書内重要項目抽出装置10300の動作アルゴリズムの一例を示す。

【0508】構造化文書要求受信部10303は、ユーザからのリクエストを受信する（ステップS11001）。リクエストを受信した構造化文書要求受信部10303は、ユーザ認証部10302を起動し、リクエストを出したユーザを認証する（ステップS11002）。構造化文書要求受信部10303は、認証されたユーザ情報とユーザのリクエストを構造化文書送受信部10306に伝える（ステップS11003）。

【0509】構造化文書送受信部10306は、外部に対し、ユーザが要求した構造化文書を要求する（ステップS11004）。外部から要求した構造化文書が送信されてきたら、構造化文書送受信部10306は、受信した構造化文書とユーザ情報を重要項目抽出部10305に渡す（ステップS11005）。

【0510】重要項目抽出部10305は、受信した構造化文書に重要項目が含まれているかどうかを検索する（ステップS11006）。重要項目が含まれている場合には、重要項目を抽出し（ステップS11008）、重要項目格納部10307を起動し、そのユーザに対応した場所に抽出した重要項目を格納する（ステップS11009）。重要項目が含まれていなかった場合には、何もしない。重要項目抽出部10305は、構造化文書送受信部10304に、受信した構造化文書を渡す（ステップS11010）。

【0511】構造化文書送受信部10304は、重要項目抽出部より渡された構造化文書をユーザに送信する（ステップS11011）。

【0512】次に、ユーザが格納されている重要項目に対してアクセス要求を出したときの文書内重要項目抽出装置10300の動作について説明する。

【0513】図42に、ユーザが格納されている重要項目に対してアクセス要求を出したときの文書内重要項目

抽出装置10300の動作アルゴリズムの一例を示す。

【0514】構造化文書要求受信部10303が、ユーザからの重要項目に対するアクセス要求を受信する（ステップS11101）。構造化文書要求受信部10303は、ユーザ認証部10302を起動し、要求を出したユーザを認証する（ステップS11102）。構造化文書要求受信部10303は、重要項目抽出部10305に対し、ユーザ情報と、アクセス要求を伝える（ステップS11103）。

【0515】重要項目抽出部10305は、アクセス許可判定部10308を起動し、ユーザが要求した重要項目へのアクセスを許可する、許可しないを判定する（ステップS11104）。アクセスが許可された場合には、重要項目抽出部10305は、重要項目格納部10307を起動し、該当する重要項目を取り出す（ステップS11106）。重要項目抽出部10305は、取り出した重要項目を構造化文書送信部10304に渡す（ステップS11108）。なお、この取り出した重要項目も構造化文書である。アクセスが許可されなかった場合には、重要項目へのアクセスが不許可であることを示す構造化文書を構造化文書送信部10304に渡す（ステップS11107）。

【0516】構造化文書送信部10304は、渡されてきた構造化文書（取り出された重要項目またはアクセスが不許可であることを示すもの）をユーザに送信する（ステップS11109）。なお、上記では、アクセスが不許可であることを示す情報を構造化文書としてユーザに渡すようにしたが、アクセスが不許可であることを示す情報を構造化文書でない情報、例えば通常のメッセージ、としてユーザに渡すようにしてもよい。

【0517】なお、実施形態3-3と同様に、重要項目に対するアクセス要求を、閲覧のみか、編集も行うか、など細かく規定し、ある重要項目に対して、そのアクセス許可リストの〈所有者〉に登録されているユーザのみがこれを編集することを可能とし、〈許可ユーザ〉に登録されているユーザは閲覧のみ可能とするなどの、細かいアクセス制御を行うようにしてもよい。

【0518】また、図37の他にも、アクセス許可リストの形式には、種々のバリエーションが考えられる。

【0519】文書内重要項目抽出装置10300は、例えば電子計算機と同じように、CPUとメモリ装置と通信装置から構成し、文書内重要項目抽出装置10300を構成する各部分10301～10308の動作を記述したプログラムをCPUで実行することにより、その機能を得ることが出来る。

【0520】また、文書内重要項目抽出装置10300を構成する各部分10301～10308は、それぞれ、もしくはその一部を、文書内重要項目抽出装置10300と同様に、CPUとメモリ装置と通信装置から構成し、それぞれの動作を記述したプログラムをCPUで

実行することにより、それぞれの機能が得られ、その集合体として構成することも可能である。

【0521】以上のように、本実施形態によれば、予め、文書内重要項目抽出装置を利用するユーザ毎に、ユーザが重要と考えるキーワードや構造化文書のタグを指定しておくことにより、文書内重要項目抽出装置が、送られてきた構造化文書から、ユーザ毎にユーザに代わって重要項目を抽出することが可能となり、ユーザの手間を大きく削減することが可能となる。また、抽出した重要項目に対して、アクセス制御情報を付加することにより、必要にして十分な重要項目をユーザ間で共有するのが容易になる。

【0522】(実施形態3-5)図43に、本実施形態(3-5)に係る文書内重要項目抽出装置の構成例を示す。

【0523】本実施形態の文書内重要項目抽出装置10700は、ユーザインタフェース部10701、ユーザ認証部10702、構造化文書要求受信部10703、構造化文書送信部10704、重要項目抽出部10705、構造化文書送受信部10706、重要項目格納部10707、アクセス許可判定部10708を備えている。

【0524】ユーザインタフェース部10701は、ユーザとのインタフェースを司る部分であり、ユーザとの情報の入出力を可能とするため、キーボード、マウスなどの入力装置、ディスプレイ、スピーカーなどの出力装置を備えている。

【0525】ユーザ認証部10702は、文書内重要項目抽出装置10700にアクセスしてくるユーザを認証する部分である。ユーザを認証する手段としては、例えば、パスワードを用いる認証方式などが利用できる。

【0526】構造化文書要求受信部10703は、ユーザからの要求を、構造化文書送受信部10706にユーザ情報とともに渡す部分である。また、ユーザが重要項目のアクセス許可情報の追加、削除、変更を要求してきた場合には、その要求をユーザ情報とともにアクセス許可判定部10708に渡し、アクセス許可情報の更新を行う。

【0527】構造化文書送受信部10706は、構造化文書要求受信部10703から渡されるユーザの要求に従って、必要となる構造化文書を外部に対して要求し、該当する構造化文書を受信する部分である。

【0528】重要項目抽出部10705は、構造化文書送受信部10706から、受信した構造化文書と、該構造化文書を要求したユーザ情報を受けとり、構造化文書内の重要項目を抽出し、重要項目格納部10707に格納するとともに、構造化文書を構造化文書送信部10704へ送る部分である。また、構造化文書要求受信部10703より、重要項目の要求があった場合には、重要項目格納部10707より要求された該重要項目を構造化

文書送信部10704へ送る。

【0529】重要項目の抽出は、予めユーザが指定したキーワードや構造化文書に附属しているタグを基に行う方法が考えられる。例えば、構造化文書の中に、ユーザが指定したタグが含まれている部分、およびキーワードが含まれている部分が存在した場合には、その該当する部分を抽出する。あるいは、これらのタグ、キーワードが含まれている構造化文書全体を対象としても良い。抽出した情報は、後に検索が可能のように、重要項目だけでなく、構造化文書の作成された日時、構造化文書を受信した日時、どこから構造化文書を受信したかを示す場所に関する情報なども合わせて記憶しても良い。

【0530】重要項目格納部10707は、重要項目抽出部10705の指示に従い、各ユーザ毎に抽出した重要項目とそれに関連する情報を格納し、読み出す部分である。重要項目格納部10707には、それを記録するための媒体が附属している。媒体としては、磁気ディスク、光ディスク、光磁気ディスク、相変化型光ディスク、CD-R、DVD-RAM、磁気テープなどが考えられる。

【0531】構造化文書送信部10704は、重要項目抽出部10705より送られてきた構造化文書を、ユーザへ送信する部分である。

【0532】アクセス許可判定部10708は、重要項目抽出部10705より、重要項目を読み出す際のアクセス許可を判断する部分である。アクセス許可判定部10708は、重要項目抽出部10705より、ある重要項目の読み出し要求と、読み出し要求を出したユーザ情報を与えられると、内部に持つアクセス許可リストに従い、読み出しの可、不可を判断し、判断結果を重要項目抽出部10705に返す。

【0533】なお、抽出された重要項目と、それに対応するアクセス許可リストの格納例は、実施形態3-3(図37)に示す例と同様であるので、ここでの説明は省略する。

【0534】次に、ユーザが外部に対して構造化文書のリクエストを出したときの文書内重要項目抽出装置10700の動作について説明する。

【0535】図44に、ユーザが外部に対して構造化文書のリクエストを出したときの文書内重要項目抽出装置10700の動作アルゴリズムの一例を示す。

【0536】ユーザインタフェース部10701がユーザからのリクエストを受け付ける(ステップS11701)。ユーザインタフェース部10701は、受け付けたリクエストを構造化文書要求受信部10703に渡す(ステップS11702)。

【0537】構造化文書要求受信部10703は、ユーザ認証部10702を起動し、リクエストを出したユーザをユーザインタフェース部10701を起動して認証する(ステップS11703)。構造化文書要求受信部

10703は、構造化文書送受信部10706に対し、ユーザ情報とユーザのリクエストを伝える（ステップS11704）。

【0538】構造化文書送受信部10706は、外部に対し、ユーザが要求した構造化文書を要求する（ステップS11705）。外部から要求した構造化文書が送信されてきたら、構造化文書送受信部10706は、受信した構造化文書とユーザ情報を重要項目抽出部10705に渡す（ステップS11706）。

【0539】重要項目抽出部10705は、受信した構造化文書に重要項目が含まれているかどうかを検索する（ステップS11707）。重要項目が含まれている場合には、重要項目を抽出し（ステップS11709）、重要項目格納部10707を起動し、そのユーザに対応した場所に抽出した重要項目を格納する（ステップS11710）。重要項目が含まれていなかった場合には、何もしない。重要項目抽出部10705は、構造化文書送信部10704に、受信した構造化文書を渡す（ステップS11711）。

【0540】構造化文書送信部10704は、ユーザインタフェース部10701に受信した構造化文書を渡し（ステップS11712）、ユーザインタフェース部10701は、受信した構造化文書をユーザに渡す（ステップS11713）。

【0541】次に、ユーザが格納されている重要項目に対してアクセス要求を出したときの文書内重要項目抽出装置10700の動作について説明する。

【0542】図45に、ユーザが格納されている重要項目に対してアクセス要求を出したときの文書内重要項目抽出装置10700の動作アルゴリズムの一例を示す。

【0543】ユーザインタフェース部10701がユーザからのリクエストを受け付ける（ステップS11801）。ユーザインタフェース部10701は、受け付けたリクエストを構造化文書要求受信部10703に渡す（ステップS11802）。

【0544】構造化文書要求受信部10703は、ユーザ認証部10702を起動し、要求を出したユーザを認証する（ステップS11803）。構造化文書要求受信部10703は、重要項目抽出部10705に対し、ユーザ情報と、アクセス要求を伝える（ステップS11804）。

【0545】重要項目抽出部10705は、アクセス許可判定部10708を起動し、ユーザが要求した重要項目へのアクセスを許可する、許可しないを判定する（ステップS11805）。アクセスが許可された場合には、重要項目抽出部10705は、重要項目格納部10707を起動し、該当する重要項目を取り出す（ステップS11807）。重要項目抽出部10705は、取り出した重要項目を構造化文書送信部10704に渡す（ステップS11809）。なお、この取り出した重要

項目も構造化文書である。アクセスが許可されなかった場合には、重要項目へのアクセスが不許可であることを示す構造化文書を構造化文書送信部10704に渡す（ステップS11808）。

【0546】構造化文書送信部10704は、渡されてきた構造化文書（取り出された重要項目またはアクセスが不許可であることを示すもの）をユーザインタフェース部10701に渡す（ステップS11810）。ユーザインタフェース部10701は、渡されてきた構造化文書をユーザに渡す（ステップS11811）。なお、上記では、アクセスが不許可であることを示す情報を構造化文書としてユーザに渡すようにしたが、アクセスが不許可であることを示す情報を構造化文書でない情報、例えば通常のメッセージ、としてユーザに渡すようにしてもよい。

【0547】なお、実施形態3-3、3-4と同様に、重要項目に対するアクセス要求を、閲覧のみか、編集も行うか、など細かく規定し、ある重要項目に対して、そのアクセス許可リストの<所有者>に登録されているユーザのみがこれを編集することを可能とし、<許可ユーザ>に登録されているユーザは閲覧のみ可能とするなどの、細かいアクセス制御を行うようにしてもよい。

【0548】次に、ユーザが重要項目に関するアクセス許可情報の追加、削除、変更を要求してきたときの文書内重要項目抽出装置10700の動作について説明する。

【0549】図46に、ユーザが重要項目に関するアクセス許可情報の追加、削除、変更を要求してきたときの文書内重要項目抽出装置10700の動作アルゴリズムの一例を示す。

【0550】ユーザインタフェース部10701がユーザからの要求を受信する（ステップS11901）。ユーザインタフェース部10701は、受け付けた要求を構造化文書要求受信部10703に渡す（ステップS11902）。

【0551】構造化文書要求受信部10703は、ユーザ認証部10702を起動し、要求を出したユーザを認証する（ステップS11903）。構造化文書受信部10703は、アクセス許可判定部10708に、ユーザ情報とユーザからの要求である、対象となる重要項目とそれに対するアクセス許可情報の変更を伝える（ステップS11904）。

【0552】アクセス許可判定部10708は、このユーザにアクセス許可情報を変更する権限があるかどうかを判断する（ステップS11905）。変更する権限がある場合には、アクセス許可情報を更新し（ステップS11906）、アクセス許可情報の変更に成功したことを示す構造化文書を構造化文書要求受信部10703に返す（ステップS11907）。変更する権限がない場合には、アクセス許可情報を変更する権限がないことを

示す構造化文書を構造化文書要求受信部10703に返す(ステップS11908)。

【0553】構造化文書要求受信部10703は、構造化文書送信部10704に、受け取った構造化文書(アクセス許可情報の変更に成功したことを示すもの、またはアクセス許可情報を変更する権限がないことを示すもの)を渡す(ステップS11909)。構造化文書送信部10704は、受け取った構造化文書をユーザインタフェース部10701に渡す(ステップS11910)。ユーザインタフェース部10701は、渡されてきた構造化文書をユーザに渡す(ステップS11911)。なお、上記では、アクセス許可情報の変更に成功したことを示す情報やアクセス許可情報を変更する権限がないことを示す情報を構造化文書としてユーザに渡すようにしたが、それら情報を構造化文書でない情報、例えば通常のメッセージ、としてユーザに渡すようにしてもよい。

【0554】なお、アクセス許可リストを変更する権限は、そのアクセス許可リストの<所有者>に登録されているユーザのみがこれを有するようにしてもよいし、そのアクセス許可リストの<所有者>と<許可ユーザ>に登録されているユーザがこれを有するようにしてもよい。

【0555】また、アクセス許可リストの変更については、そのアクセス許可リストの<許可ユーザ>にのみ他のユーザを追加できるようにしてもよいし、さらに<所有者>にも他のユーザを追加できるようにしてもよい。

【0556】また、アクセス許可リストには、<所有者>と<許可ユーザ>の区別を設けず、その重要項目を得るもととなった構造化文書の要求を出したユーザと、その後許可されたユーザとをまとめて登録するようにしてもよい。この場合には、アクセス許可リストに登録された全ユーザに変更権限を与えて構わない。

【0557】また、図37の他にも、アクセス許可リストの形式には、種々のバリエーションが考えられる。

【0558】文書内重要項目抽出装置10700は、例えば電子計算機と同じように、CPUとメモリ装置と通信装置から構成し、文書内重要項目抽出装置10700を構成する各部分10701～10708の動作を記述したプログラムをCPUで実行することにより、その機能を得ることが出来る。

【0559】また、文書内重要項目抽出装置10700を構成する各部分10701～10708は、それぞれ、もしくはその一部を文書内重要項目抽出装置10700と同様に、CPUとメモリ装置と通信装置から構成し、それぞれの動作を記述したプログラムをCPUで実行することにより、それぞれの機能が得られ、その集合体として構成することも可能である。

【0560】以上のように、本実施形態によれば、予め、文書内重要項目抽出装置を利用するユーザ毎に、ユ

ーザが重要と考えるキーワードや構造化文書のタグを指定しておくことにより、文書内重要項目抽出装置が、送られてきた構造化文書から、ユーザ毎にユーザに代わって重要項目を抽出することが可能となり、ユーザの手間を大きく削減することが可能となる。また、抽出した重要項目に対して、アクセス許可情報を付加することにより、必要にして十分な重要項目をユーザ間で共有するのが容易になる。また、ユーザの要求に応じ、ユーザが保存されている重要項目のアクセス許可情報を追加、削除、変更する機能を持たせることで、異なる重要項目を、ユーザ間で容易に適切に共有することが可能となり、重要項目の共有にかかるユーザの手間を大きく削減することが可能となる。また、重要項目を共有することで、文書内重要項目抽出装置が具備する記憶装置の効率的な利用が可能となる。

【0561】(実施形態3-6)図47に、本実施形態(3-6)に係る文書内重要項目抽出装置の構成例を示す。

【0562】本実施形態の文書内重要項目抽出装置10600は、ユーザ認証部10602、構造化文書要求受信部10603、構造化文書送信部10604、重要項目抽出部10605、構造化文書送受信部10606、重要項目格納部10607、アクセス許可判定部10608を備えている。

【0563】ユーザ認証部10602は、文書内重要項目抽出装置10600に、ブラウザなどを用いてアクセスしてくるユーザ10601を認証する部分である。ユーザを認証する手段としては、例えば、パスワードを用いる認証方式などが利用できる。

【0564】構造化文書要求受信部10603は、ユーザ10601からの要求を、構造化文書送受信部10606にユーザ情報とともに渡す部分である。また、ユーザ10601が重要項目のアクセス許可情報の追加、削除、変更を要求してきた場合には、その要求をユーザ情報とともにアクセス許可判定部10608に渡し、アクセス許可情報の変更を行う。

【0565】構造化文書送受信部10606は、構造化文書要求受信部10603から渡されるユーザ10601の要求に従って、必要となる構造化文書を外部に対して要求し、該当する構造化文書を受信する部分である。

【0566】重要項目抽出部10605は、構造化文書送受信部10606から、受信した構造化文書と、該構造化文書を要求したユーザ情報を受けとり、構造化文書内の重要項目を抽出し、重要項目格納部10607に格納するとともに、構造化文書を構造化文書送信部10604へ送る部分である。また、構造化文書要求受信部10603より、重要項目の要求があった場合には、重要項目格納部10607より要求された該重要項目を構造化文書送信部10604へ送る。

【0567】重要項目の抽出は、予めユーザ10601



が指定したキーワードや構造化文書に附属しているタグを基に行う方法が考えられる。例えば、構造化文書の中に、ユーザ10601が指定したタグが含まれている部分、およびキーワードが含まれている部分が存在した場合には、その該当する部分を抽出する。あるいは、これらのタグ、キーワードが含まれている構造化文書全体を対象としても良い。抽出した情報は、後に検索が可能なように、重要項目だけでなく、構造化文書の作成された日時、構造化文書を受信した日時、どこから構造化文書を受信したかを示す場所に関する情報なども合わせて記憶する。

【0568】重要項目格納部10607は、重要項目抽出部10605の指示に従い、各ユーザ毎に抽出した重要項目とそれに関連する情報を格納し、読み出す部分である。重要項目格納部10607には、それを記録するための媒体が附属している。媒体としては、磁気ディスク、光ディスク、光磁気ディスク、相変化型光ディスク、CD-R、DVD-RAM、磁気テープなどが考えられる。

【0569】構造化文書送信部10604は、重要項目抽出部10605より送られてきた構造化文書を、ユーザ10601へ送信する部分である。

【0570】アクセス許可判定部10608は、重要項目抽出部10605より、重要項目を読み出す際のアクセス許可を判断する部分である。アクセス許可判定部10608は、重要項目抽出部10605より、ある重要項目の読み出し要求と、読み出し要求を出したユーザ情報を与えられると、内部に持つアクセス許可リストに従い、読み出しの可、不可を判断し、判断結果を重要項目抽出部10605に返す。

【0571】なお、抽出された重要項目と、それに対応するアクセス許可リストの格納例は、実施形態3-3(図37)に示す例と同様であるので、ここでの説明は省略する。

【0572】次に、ユーザが外部に対して構造化文書のリクエストを出したときの文書内重要項目抽出装置10600の動作について説明する。

【0573】図48に、ユーザが外部に対して構造化文書のリクエストを出したときの文書内重要項目抽出装置10600の動作アルゴリズムの一例を示す。

【0574】構造化文書要求受信部10603は、ユーザからのリクエストを受信する(ステップS11401)。リクエストを受信した構造化文書要求受信部10603は、ユーザ認証部10602を起動し、リクエストを出したユーザを認証する(ステップS11402)。構造化文書要求受信部10603は、認証されたユーザ情報とユーザのリクエストを構造化文書受信部10606に伝える(ステップS11403)。

【0575】構造化文書受信部10606は、外部に対し、ユーザが要求した構造化文書を要求する(ステッ

プS11404)。外部から要求した構造化文書が送信されてきたら、構造化文書受信部10606は、受信した構造化文書とユーザ情報を重要項目抽出部10605に渡す(ステップS11405)。

【0576】重要項目抽出部10605は、受信した構造化文書に重要項目が含まれているかどうかを検索する(ステップS11406)。重要項目が含まれている場合には、重要項目を抽出し(ステップS11408)、重要項目格納部10607を起動し、そのユーザに対応した場所に抽出した重要項目を格納する(ステップS11409)。重要項目が含まれていなかった場合には、何もしない。重要項目抽出部10605は、構造化文書送信部10604に、受信した構造化文書を渡す(ステップS11410)。

【0577】構造化文書送信部10604は、重要項目抽出部より渡された構造化文書をユーザに送信する(ステップS11411)。

【0578】次に、ユーザが格納されている重要項目に対してアクセス要求を出したときの文書内重要項目抽出装置10600の動作について説明する。

【0579】図49に、ユーザが格納されている重要項目に対してアクセス要求を出したときの文書内重要項目抽出装置10600の動作アルゴリズムの一例を示す。

【0580】構造化文書要求受信部10603が、ユーザからの重要項目に対するアクセス要求を受信する(ステップS11501)。構造化文書要求受信部10603は、ユーザ認証部10602を起動し、要求を出したユーザを認証する(ステップS11502)。構造化文書要求受信部10603は、重要項目抽出部10605に対し、ユーザ情報と、アクセス要求を伝える(ステップS11503)。

【0581】重要項目抽出部10605は、アクセス許可判定部10608を起動し、ユーザが要求した重要項目へのアクセスを許可する、許可しないを判定する(ステップS11504)。アクセスが許可された場合には、重要項目抽出部10605は、重要項目格納部10607を起動し、該当する重要項目を取り出す(ステップS11506)。重要項目抽出部10605は、取り出した重要項目を構造化文書送信部10604に渡す(ステップS11508)。なお、この取り出した重要項目も構造化文書である。アクセスが許可されなかった場合には、重要項目へのアクセスが不許可であることを示す構造化文書を構造化文書送信部10604に渡す(ステップS11507)。

【0582】構造化文書送信部10604は、渡されてきた構造化文書(取り出された重要項目またはアクセスが不許可であることを示すもの)をユーザに送信する(ステップS11509)。なお、上記では、アクセスが不許可であることを示す情報を構造化文書としてユーザに渡すようにしたが、アクセスが不許可であることを

示す情報を構造化文書でない情報、例えば通常のメッセージ、としてユーザに渡すようにしてもよい。

【0583】なお、実施形態3-3~3-5と同様に、重要項目に対するアクセス要求を、閲覧のみか、編集も行うか、など細かく規定し、ある重要項目に対して、そのアクセス許可リストの<所有者>に登録されているユーザのみがこれを編集することを可能とし、<許可ユーザ>に登録されているユーザは閲覧のみ可能とするなどの、細かいアクセス制御を行うようにしてもよい。

【0584】次に、ユーザが重要項目に関するアクセス許可情報の追加、削除、変更を要求してきたときの文書内重要項目抽出装置10600の動作について説明する。

【0585】図50に、ユーザが重要項目に関するアクセス許可情報の追加、削除、変更を要求してきたときの文書内重要項目抽出装置10600の動作アルゴリズムの一例を示す。

【0586】構造化文書受信部10603がユーザからの要求を受信する(ステップS11601)。構造化文書受信部10603は、ユーザ認証部10602を起動し、要求を出したユーザを認証する(ステップS11602)。構造化文書受信部10603は、アクセス許可判定部10608に、ユーザ情報とユーザからの要求である、対象となる重要項目とそれに対するアクセス許可情報の変更を伝える(ステップS11603)。

【0587】アクセス許可判定部10608は、このユーザにアクセス許可情報を変更する権限があるかどうかを判断する(ステップS11604)。変更する権限がある場合には、アクセス許可情報を更新し(ステップS11605)、アクセス許可情報の変更に成功したことを示す構造化文書を構造化文書要求受信部10603に返す(ステップS11606)。変更する権限がない場合には、アクセス許可情報を変更する権限がないことを示す構造化文書を構造化文書要求受信部10603に返す(ステップS11607)。

【0588】構造化文書要求受信部10603は、構造化文書送信部10604に、受け取った構造化文書(アクセス許可情報の変更に成功したことを示すもの、またはアクセス許可情報を変更する権限がないことを示すもの)を渡す(ステップS11608)。構造化文書送信部10604は、受け取った構造化文書をユーザに送信する(ステップS11609)。なお、上記では、アクセス許可情報の変更に成功したことを示す情報やアクセス許可情報を変更する権限がないことを示す情報を構造化文書としてユーザに渡すようにしたが、それら情報を構造化文書でない情報、例えば通常のメッセージ、としてユーザに渡すようにしてもよい。

【0589】なお、実施形態3-5と同様に、アクセス許可リストを変更する権限は、そのアクセス許可リストの<所有者>に登録されているユーザのみがこれを有す

るようにしてもよいし、そのアクセス許可リストの<所有者>と<許可ユーザ>に登録されているユーザがこれを有するようにしてもよい。

【0590】また、アクセス許可リストの変更については、そのアクセス許可リストの<許可ユーザ>にのみ他のユーザを追加できるようにしてもよいし、さらに<所有者>にも他のユーザを追加できるようにしてもよい。

【0591】また、アクセス許可リストには、<所有者>と<許可ユーザ>の区別を設けず、その重要項目を得るもととなった構造化文書の要求を出したユーザと、その後に許可されたユーザとをまとめて登録するようにしてもよい。この場合には、アクセス許可リストに登録された全ユーザに変更権限を与えて構わない。

【0592】また、図37の他にも、アクセス許可リストの形式には、種々のバリエーションが考えられる。

【0593】文書内重要項目抽出装置10600は、例えば電子計算機と同じように、CPUとメモリ装置と通信装置から構成し、文書内重要項目抽出装置10600を構成する各部分10601~10608の動作を記述したプログラムをCPUで実行することにより、その機能を得ることが出来る。

【0594】また、文書内重要項目抽出装置10600を構成する各部分10601~10608は、それぞれ、もしくはその一部を文書内重要項目抽出装置10600と同様に、CPUとメモリ装置と通信装置から構成し、それぞれの動作を記述したプログラムをCPUで実行することにより、それぞれの機能が得られ、その集合体として構成することも可能である。

【0595】以上のように、本実施形態によれば、予め、文書内重要項目抽出装置を利用するユーザ毎に、ユーザが重要と考えるキーワードや構造化文書のタグを指定しておくことにより、文書内重要項目抽出装置が、送られてきた構造化文書から、ユーザ毎にユーザに代わって重要項目を抽出することが可能となり、ユーザの手間を大きく削減することが可能となる。また、抽出した重要項目に対して、アクセス許可情報を付加することにより、必要にして十分な重要項目をユーザ間で共有するのが容易になる。また、ユーザの要求に応じ、ユーザが保存されている重要項目のアクセス許可情報を追加、削除、変更する機能を持たせることで、異なる重要項目を、ユーザ間で容易に適切に共有することが可能となり、重要項目の共有にかかるユーザの手間を大きく削減することが可能となる。また、重要項目を共有することで、文書内重要項目抽出装置が具備する記憶装置の効率的な利用が可能となる。

【0596】(実施形態3-7)図51に、本実施形態(3-7)に係る文書内重要項目抽出装置の構成例を示す。

【0597】本実施形態に係る文書内重要項目抽出装置

は、例えば実施形態 3-5、3-6 のような構成において、複数のユーザが同一の重要項目抽出条件を設定しており（もしくは設定した重要項目抽出条件の少なくとも一部に共通するものがあり）、かつ該複数のユーザが同一の構造化文書を要求した場合に、ユーザ間で共通の重要項目が抽出されることから、受信した構造化文書から抽出された重要項目のうちユーザ間で共通の重要項目についてはこれを共有できる形態で保存するようにしたのである。

【0598】なお、このような文書内重要項目抽出装置として、実施形態 3-5、3-6 のように、データ送受信装置に組み込む形態と、データ中継装置に組み込む形態とが考えられるが、両者の違いは、ユーザインタフェース部を持つか否かのみであり、本実施形態に係る抽出した共通の重要項目をユーザ間で共有可能な形態で保存する機能に関しては同様であるため、本実施形態では文書内重要項目抽出装置をデータ中継装置に組み込む場合については詳細な説明は省略する。

【0599】さて、図 51 に示されるように、本実施形態の文書内重要項目抽出装置 13000 は、ユーザ認証部 13002、構造化文書要求部 13003、構造化文書送信部 13004、重要項目抽出部 13005、構造化文書送受信部 13006、重要項目格納部 13007、アクセス許可判定部 13008 を備えている。また、重要項目抽出部 13005 は、実施形態 3-6 のように重要項目の抽出や（共通でない）重要項目の保存などに関する処理の機能を有するのに加えて、共通の重要項目に関する処理を行うための共通重要項目処理部 13051 を有する。

【0600】ユーザ認証部 13002 は、文書内重要項目抽出装置 13000 に、ブラウザなどを用いてアクセスしてくるユーザ 13001 を認証する部分である。ユーザを認証する手段としては、例えば、パスワードを用いる認証方式などが利用できる。

【0601】構造化文書要求受信部 13003 は、ユーザ 13001 からの要求を、構造化文書送受信部 13006 にユーザ情報とともに渡す部分である。また、ユーザ 13001 が重要項目のアクセス許可情報の追加、削除、変更を要求してきた場合には、その要求をユーザ情報とともにアクセス許可判定部 13008 に渡し、アクセス許可情報の更新を行う。

【0602】構造化文書送受信部 13006 は、構造化文書要求受信部 13003 から渡されるユーザ 13001 の要求に従って、必要となる構造化文書を外部に対して要求し、該当する構造化文書を受信する部分である。

【0603】重要項目抽出部 13005 は、構造化文書送受信部 13006 から、受信した構造化文書と、該構造化文書を要求したユーザ情報を受け取り、構造化文書内の重要項目を抽出し、重要項目格納部 13007 に格

納するとともに、構造化文書を構造化文書送信部 13004 へ送る部分である。また、構造化文書要求受信部 13003 より、重要項目の要求があった場合には、重要項目格納部 13007 より要求された該重要項目を構造化文書送信部 13004 へ送る。さらに、共通重要項目処理を行うための部分である（共通重要項目処理部 13051）。

【0604】重要項目の抽出は、予めユーザ 13001 が指定したキーワードや、構造化文書に付属しているタグを基に行う方法が考えられる。例えば、構造化文書の中に、ユーザ 13001 が指定したタグが含まれている部分、およびキーワードが含まれている部分が存在した場合には、その該当する部分を抽出する。あるいは、これらのタグ、キーワードが含まれている構造化文書全体を対象としてもよい。抽出した情報は、後に検索が可能ないように、重要項目だけでなく、構造化文書の作成された日時、構造化文書を受信した日時、どこから構造化文書を受信したかを示す場所に関する情報なども併せて記憶する。

【0605】なお、本実施形態では、共通重要項目処理によって、過去に同一構造化文書から同一重要項目が抽出されている場合には、今回の要求によって抽出された重要項目を別途保存することはせずに、当該保存されている該当重要項目に付随する情報（本実施形態例では、アクセス許可リストとする）に必要な修正をかけるようにしている（修正が不要なこともある）。

【0606】重要項目格納部 13007 は、重要項目抽出部 13005 の指示に従い、各ユーザ毎に抽出した重要項目とそれに関連する情報を格納し、読み出す部分である。重要項目格納部 13007 には、それを記録するための媒体が付属している。媒体としては、磁気ディスク、光ディスク、光磁気ディスク、相変化型光ディスク、CD-R、DVD-RAM、磁気テープなどが考えられる。

【0607】構造化文書送信部 13004 は、重要項目抽出部 13005 より送られてきた構造化文書を、ユーザ 13001 へ送信する部分である。

【0608】アクセス許可判定部 13008 は、重要項目抽出部 13005 より、重要項目を読み出す際のアクセス許可を判断する部分である。アクセス許可判定部 13008 は、重要項目抽出部 13005 より、ある重要項目の読み出し要求と、読み出し要求を出したユーザ情報を与えられると、内部に持つアクセス許可リストに従い、読み出しの可、不可を判断し、判断結果を重要項目抽出部 13005 に返す。

【0609】なお、抽出された重要項目と、それに対応するアクセス許可リストの形式としては、実施形態 3-3（図 37）に示す例を用いることができるが、その他にも種々の形式のアクセス許可リストを用いることができ、本実施形態では、いくつかの形式のアクセス許可リ

ストについて説明する。

【0610】次に、ユーザが外部に対して構造化文書のリクエストを出したときの文書内重要項目抽出装置13000の動作について説明する。

【0611】図52に、ユーザが外部に対して構造化文書のリクエストを出したときの文書内重要項目抽出装置13000の動作アルゴリズムの一例を示す。

【0612】構造化文書要求受信部13003は、ユーザからのリクエストを受信する(ステップS13101)。リクエストを受信した構造化文書要求受信部13003は、ユーザ認証部13002を起動し、リクエストを出したユーザを認証する(ステップS13102)。構造化文書要求受信部13003は、認証されたユーザ情報とユーザのリクエストを構造化文書送受信部13006に伝える(ステップS13103)。

【0613】構造化文書送受信部13006は、外部に対し、ユーザが要求した構造化文書を要求する(ステップS13104)。外部から要求した構造化文書が送信されてきたら、構造化文書送受信部13006は、受信した構造化文書とユーザ情報を重要項目抽出部13005に渡す(ステップS13105)。

【0614】重要項目抽出部13005は、受信した構造化文書に重要項目が含まれているかどうかを検索する(ステップS13106)。

【0615】重要項目が含まれている場合には、重要項目を抽出する(ステップS13108)。ここで、本実施形態では、すぐに抽出された重要項目を保存せずに、共通重要項目処理部13051により、共通重要項目処理を行う(ステップS13109)。ここでは、詳しくは後述するが、当該重要項目が過去に同一構造化文書から抽出され保存されているか否かを調べたり、保存されている場合のアクセス許可リストの更新の仕方またはそれを更新しないことの決定などを行う。次に、重要項目格納部13007を起動し、そのユーザに対応した場所に抽出した重要項目を格納(保存されていない場合)するか、または該当する重要項目について必要に応じてアクセス許可リストを更新(保存されている場合)する(ステップS13110)。

【0616】一方、受信した構造化文書に重要項目が含まれていなかった場合には、何もしない。

【0617】重要項目抽出部13005は、構造化文書送信部13004に、受信した構造化文書を渡す(ステップS13111)。

【0618】構造化文書送信部13004は、重要項目抽出部13005より渡された構造化文書をユーザに送信する(ステップS13112)。

【0619】次に、重要項目を抽出するときの重要項目抽出部13005の動作について説明する。

【0620】図53に、重要項目抽出部13005の重要項目を抽出するときのアルゴリズムの一例を示す。な

お、本アルゴリズムでは、複数の要求とこれに対する応答を同時に扱うことはせず、個々の要求とその応答がユーザ毎に独立であるものとして動作することを仮定している。また、抽出した重要項目は、図54のような形式で格納されているものとする。

【0621】まず、構造化文書送受信部13006から構造化文書を受け取る(ステップS13201)。

【0622】次に、ユーザ認証部13002より、この構造化文書を要求したユーザ情報を得る(ステップS13202)。

【0623】次に、設定されているタグ、キーワードより、構造化文書から、重要項目を抽出する(ステップS13203)。

【0624】ここで、抽出した重要項目を、過去に同一の構造化文書から抽出したことがあるかどうかを調べる(ステップS13204)。

【0625】もし過去に同一の構造化文書から同一の重要項目を抽出したことがある場合には、過去に抽出したときの該当重要項目のアクセス許可リストを調べる(ステップS13205)。すなわち、この構造化文書を要求したユーザのユーザ名と、該当重要項目のアクセス許可リストに含まれるユーザ名とを比較する。当該要求ユーザのユーザ名が当該アクセス許可リストに含まれていない場合には(ステップS13206)、当該アクセス許可リストに当該要求ユーザのユーザ名を加えるなどの必要な処理を行う(ステップS13207)。なお、このアクセス許可リストの更新の形態にはアクセス許可リストの形式などに応じて種々のバリエーションがある。当該アクセス許可リストに当該要求ユーザのユーザ名が含まれる場合には(ステップS13206)、既に該当重要項目にアクセス可能であるため、処理を終える。

【0626】一方、ステップS13204において、今回抽出した重要項目を、過去に同一の構造化文書から抽出したことがない場合には、基本的には実施形態3-6、3-7と同様で、木構造を構成している格納済みの重要項目に、新たに今回抽出した重要項目を付け加え(ステップS13208)、そして、この重要項目に関するアクセス許可リストに、このユーザ名を設定し(ステップS13209)、処理を終える。

【0627】なお、複数ユーザが同一の抽出条件を設定しており、かつ該複数ユーザが同一の構造化文書を受信したときの処理は、上記アルゴリズムを、ユーザ毎に、抽出する重要項目の分だけ繰り返すことにより、実現することができる。

【0628】ここで、アクセス許可リストのいくつかのバリエーションについて説明する。

【0629】まず、図37にて例示した形式のアクセス許可リストについて説明する。

【0630】この場合には、ステップS13205にて、該当重要項目のアクセス許可リストの<所有者>お

よび<許可ユーザ>を調べ、当該要求ユーザのユーザ名が<所有者>と<許可ユーザ>のいずれにも存在しない場合に、ステップS13207にて、該当重要項目のアクセス許可リストの<許可ユーザ>に当該要求ユーザのユーザ名を加える。

【0631】次に、図54にて例示する形式のアクセス許可リストについて説明する。

【0632】図54では、アクセス許可リスト内に<所有者>と<許可ユーザ>の構造を設けず、すべて所有者として<アクセス許可リスト>の直下に付加するようにしたものである。

【0633】この場合には、ステップS13205にて、該当重要項目のアクセス許可リストを調べ、当該要求ユーザのユーザ名が存在しない場合に、ステップS13207にて、該当重要項目のアクセス許可リストに当該要求ユーザのユーザ名を加える。

【0634】ところで、複数ユーザが同一の抽出した重要項目を共有しており、該複数ユーザのうちの1ユーザ、またはそれ以上のユーザが、格納されている重要項目を編集したい場合には、編集したいユーザの分の重要項目をコピーし、それを編集するようにすると好ましい。例えば、図54においては、項目2に関して、ユーザA、B、Cが共有していたが、ここで、ユーザBが重要項目を編集したい場合、図55に示すように、ユーザBを項目2のアクセス許可リストから取り除き、新たに項目3を作成して項目2の重要項目をコピーし、項目3のアクセス許可リストにユーザBを加えればよい。

【0635】次に、さらに他の形式のアクセス許可リストの例について説明する。

【0636】上記では、抽出した同一の重要項目を共有するために、図37、図54に示すような、アクセス許可の仕組みを取り入れたが、仕組みとしては、図56、図57に示すような方法も考えられる。

【0637】図56では、抽出した同一の重要項目に対して、図37に示す例と同様に、許可ユーザのリストを追加するが、その際に所有者の部分で、最初に該重要項目を抽出したユーザだけでなく、後に同一の重要項目を抽出したユーザについても、該ユーザを所有者に付け加える方法である。これにより、同一の重要項目を抽出したユーザは、全て該重要項目の所有者となり、該重要項目に対するアクセス許可ユーザのリストを編集することができる。

【0638】図56の例は、先にユーザAがある構造化文書から重要項目を抽出した後に、ユーザBが同一の構造化文書から同一の重要項目を抽出したときの格納状態を示したものであり、変更された部分を、太線で示している。

【0639】なお、該重要項目を編集する場合には、いずれのユーザであっても、図55に示すように、重要項目のコピーを作成した上で、編集させるようにするのが

好ましい。

【0640】次に、図57は、抽出した同一の重要項目に対して、所有者の部分にユーザを追加するのではなく、所有者、および許可ユーザから構成される構造自体を増やす方法である。図57の例は、図56の例と同様、後からユーザBが同一の構造化文書から同一の重要項目を抽出したときの格納状態を示したものであり、変更された部分を太線で示している。

【0641】なお、該重要項目を編集する場合には、図55に示すように、重要項目のコピーを作成した上で、編集させるようにするのが好ましい。

【0642】なお、ユーザが、格納されている重要項目に対して、アクセス要求を出したときの、文書内重要項目抽出装置13000の動作アルゴリズム、およびユーザが重要項目に関するアクセス許可情報の追加、削除、変更を要求してきたときの文書内重要項目抽出装置13000の動作アルゴリズムは、基本的に、実施形態3-6に示す例と同様であるので、このでの詳細な説明は省略する。

【0643】また、実施形態3-5、3-6と同様に、図37、図56、図57の例の場合に、重要項目に対するアクセス要求を、閲覧のみか、編集も行うか、など細かく規定し、ある重要項目に対して、そのアクセス許可リストの<所有者>に登録されているユーザのみがこれを編集することを可能とし、<許可ユーザ>に登録されているユーザは閲覧のみ可能とするなどの、細かいアクセス制御を行うようにしてもよい。

【0644】また、実施形態3-5、3-6と同様に、図37、図56、図57の例の場合に、アクセス許可リストを変更する権限は、そのアクセス許可リストの<所有者>に登録されているユーザのみがこれを有するようにしてもよいし、そのアクセス許可リストの<所有者>と<許可ユーザ>に登録されているユーザがこれを有するようにしてもよい。

【0645】また、アクセス許可リストの変更については、そのアクセス許可リストの<許可ユーザ>にのみ他のユーザを追加できるようにしてもよいし、さらに<所有者>にも他のユーザを追加できるようにしてもよい。

【0646】また、アクセス許可リストには、<所有者>と<許可ユーザ>の区別を設けず、その重要項目を得るもとなった構造化文書の要求を出したユーザと、その後に許可されたユーザとをまとめて登録するようにしてもよい。この場合には、アクセス許可リストに登録された全ユーザに変更権限を与えて構わない。

【0647】また、図37、図54、図56、図57の他にも、アクセス許可リストの形式には、種々のバリエーションが考えられる。

【0648】文書内重要項目抽出装置13000は、例えば電子計算機と同じように、CPUとメモリ装置と通

信装置から構成し、文書内重要項目抽出装置 13000 を構成する各部分 13001～13008 の動作を記述させたプログラムを CPU で実行することにより、その機能を得ることが出来る。

【0649】また、文書内重要項目抽出装置 1300 を構成する各部分 13001～13008 は、それぞれ、もしくはその一部を文書内重要項目抽出装置 13000 と同様に、CPU とメモリ装置と通信装置から構成し、それぞれの動作を記述したプログラムを CPU で実行することにより、それぞれの機能が得られ、その集合体として構成することも可能である。

【0650】以上のように、本実施形態によれば、予め、文書内重要項目抽出装置を利用するユーザ毎に、ユーザが重要と考えるキーワードや構造化文書のタグを指定しておくことにより、文書内重要項目抽出装置が、送られてきた構造化文書から、ユーザ毎にユーザに代わって重要項目を抽出することが可能となり、ユーザの手間を大きく削減することが可能となる。また、抽出した重要項目に対して、アクセス許可情報を付加することにより、必要にして十分な重要項目をユーザ間で共有するのが容易になる。また、ユーザの要求に応じ、ユーザが保存されている重要項目のアクセス許可情報を追加、削除、変更する機能を持たせることで、異なる重要項目を、ユーザ間で容易に適切に共有することが可能となり、重要項目の共有にかかるユーザの手間を大きく削減することが可能となる。また、重要項目を共有することで、文書内重要項目抽出装置が具備する記憶装置の効率的な利用が可能となる。

【0651】以上、第3の実施形態によれば、本文書内重要項目抽出装置は、ユーザが指定した抽出すべき重要項目を自動的に構造化文書から抽出し、保存するが、抽出した重要項目をユーザ毎に分離して保存することにより、特定のユーザが保存するよう指定した重要項目の読み出し、検索が容易になる。また、保存した重要項目毎に、アクセス許可リストを持つため、保存した重要項目を複数のユーザ間で、容易に共有することが可能になり、記憶装置の効率的な利用が可能となる。

【0652】なお、第3の実施形態（実施形態 3-1～3-7 のそれぞれ）においても、第2の実施形態のような付加項目に関する機能を追加してもよい。

【0653】なお、以上の各機能は、ソフトウェアとしても実現可能である。

【0654】また、本実施形態は、コンピュータに所定の手順を実行させるための（あるいはコンピュータを所定の手段として機能させるための、あるいはコンピュータに所定の機能を実現させるための）プログラムを記録したコンピュータ読取り可能な記録媒体として実施することもできる。

【0655】本発明は、上述した実施の形態に限定されるものではなく、その技術的範囲において種々変形して

実施することができる。

【0656】

【発明の効果】本発明によれば、ユーザに負担をかけずに構造化文書のうち重要な内容を含む部分を自動的に保存することができる。

【図面の簡単な説明】

【図1】本発明の第1の実施形態に係る文書内重要項目抽出装置について説明するための図

【図2】同実施形態に係る文書内重要項目抽出装置の WWW システムにおける重要項目抽出動作手順の一例を示すフローチャート

【図3】同実施形態に係る文書内重要項目抽出装置の電子メールシステムにおける重要項目抽出動作手順の一例を示すフローチャート

【図4】同実施形態に係る文書内重要項目抽出装置の電子ニュースシステムにおける重要項目抽出動作手順の一例を示すフローチャート

【図5】同実施形態に係る文書内重要項目抽出装置における重要項目の抽出方法を設定するための構成について説明するための図

【図6】同実施形態に係る文書内重要項目抽出装置における重要項目の抽出方法の設定手順の一例を示すフローチャート

【図7】同実施形態に係る文書内重要項目抽出装置における重要項目を抽出する手順の一例を示すフローチャート

【図8】構造化文書の例を示す図

【図9】重要項目の抽出処理途中の構造化文書の状態の例を示す図

【図10】重要項目の抽出方法の設定例を示す図

【図11】重要項目の抽出時に用いる類語データベースの一例を示す図

【図12】抽出された重要項目の例を示す図

【図13】重要項目を格納する重要項目データベースの一例を示す図

【図14】重要項目を格納する重要項目データベースの他の例を示す図

【図15】同実施形態に係る構造化文書表示クライアントにおいて重要項目が含まれていない構造化文書の表示例を示す図

【図16】同実施形態に係る構造化文書表示クライアントにおいて表示された構造化文書内の重要項目に対する指定、追加、変更、削除を行うための表示画面例を示す図

【図17】同実施形態に係る構造化文書表示クライアントの重要項目指定画面の一例を示す図

【図18】同実施形態に係る構造化文書表示クライアントの重要項目追加画面の一例を示す図

【図19】同実施形態に係る構造化文書表示クライアントの重要項目変更画面の一例を示す図



【図20】同実施形態に係る構造化文書表示クライアントの重要項目削除画面の一例を示す図

【図21】本発明の第2の実施形態における重要項目および付加項目を格納する重要項目・付加項目データベースの一例を示す図

【図22】同実施形態に係る文書内重要項目抽出装置の一構成例を示す図

【図23】重要な項目が満たすべき条件を新たに構造化文書内重要項目条件テーブルに追加したい場合の手続きの一例を示すフローチャート

【図24】構造化文書から抽出した重要項目に関連する付加項目が満たすべき条件を新たに構造化文書内重要項目条件テーブルに追加したい場合の手続きの一例を示すフローチャート

【図25】ユーザが構造化文書の閲覧要求を出してから構造化文書内重要項目抽出装置に重要項目と付加項目が保存されるまでに行われる手続きの一例を示すフローチャート

【図26】文書内重要項目抽出装置内に格納されている項目の検索を行う際の手続きの一例を示すフローチャート

【図27】同実施形態に係る文書内重要項目抽出装置の他の構成例を示す図

【図28】重要な項目が満たすべき条件を新たに構造化文書内重要項目条件テーブルに追加したい場合の手続きの一例を示すフローチャート

【図29】構造化文書から抽出した重要項目に関連する付加項目が満たすべき条件を新たに構造化文書内重要項目条件テーブルに追加したい場合の手続きの一例を示すフローチャート

【図30】構造化文書の閲覧要求を出してから構造化文書内重要項目抽出装置に重要項目と付加項目が保存されるまでに行われる手続きの一例を示すフローチャート

【図31】文書内重要項目抽出装置内に格納されている項目の検索を行う際の手続きの一例を示すフローチャート

【図32】本発明の第3の実施形態に係る文書内重要項目抽出装置の構成の一例を示す図

【図33】重要項目を抽出し格納する手続きの一例を示すフローチャート

【図34】同実施形態に係る文書内重要項目抽出装置の構成の他の例を示す図

【図35】重要項目を抽出し格納する手続きの他の例を示すフローチャート

【図36】同実施形態に係る文書内重要項目抽出装置の構成のさらに他の例を示す図

【図37】重要項目と該重要項目に対応するアクセス許可リストの格納例を示す図

【図38】重要項目を抽出し格納する手続きのさらに他の例を示すフローチャート

【図39】保存されている重要項目に対するアクセス要求を処理する手続きの一例を示すフローチャート

【図40】同実施形態に係る文書内重要項目抽出装置の構成のさらに他の例を示す図

【図41】重要項目を抽出し格納する手続きのさらに他の例を示すフローチャート

【図42】保存されている重要項目に対するアクセス要求を処理する手続きの他の例を示すフローチャート

【図43】同実施形態に係る文書内重要項目抽出装置の構成のさらに他の例を示す図

【図44】重要項目を抽出し格納する手続きのさらに他の例を示すフローチャート

【図45】保存されている重要項目に対するアクセス要求を処理する手続きのさらに他の例を示すフローチャートを示す図

【図46】保存されている重要項目に対するアクセス要求を変更する手続きの一例を示すフローチャート

【図47】同実施形態に係る文書内重要項目抽出装置の構成のさらに他の例を示す図

【図48】重要項目を抽出し格納する手続きのさらに他の例を示すフローチャート

【図49】保存されている重要項目に対するアクセス要求を処理する手続きのさらに他の例を示すフローチャート

【図50】保存されている重要項目に対するアクセス要求を変更する手続きの他の例を示すフローチャート

【図51】同実施形態に係る文書内重要項目抽出装置の構成のさらに他の例を示す図

【図52】重要項目を抽出し格納する手続きのさらに他の例を示すフローチャート

【図53】重要項目を抽出し格納するときの重要項目抽出部の手続きの一例を示すフローチャート

【図54】重要項目と該重要項目に対応するアクセス許可リストの他の格納例を示す図

【図55】重要項目と該重要項目に対応するアクセス許可リストの他の格納例を示す図

【図56】重要項目と該重要項目に対応するアクセス許可リストのさらに他の格納例を示す図

【図57】重要項目と該重要項目に対応するアクセス許可リストのさらに他の格納例を示す図

【符号の説明】

101…WWWページ作成者

102…電子メール送信者

103…電子ニュース記事送信者

104…WWWサーバ

105…送信側電子メールサーバ

106…送信側電子ニュースサーバ

107…プロキシサーバ

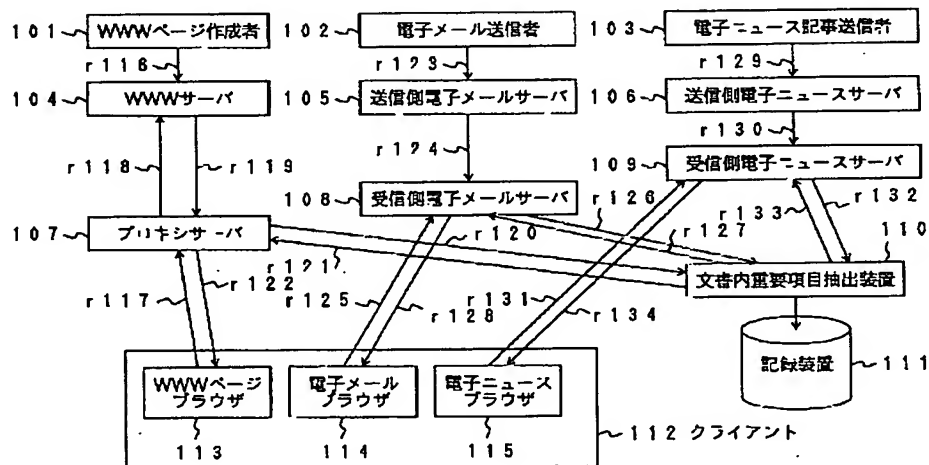
108…受信側電子メールサーバ

109…受信側電子ニュースサーバ

110, 220, 10200~10700, 13000  
 …文書内重要項目抽出装置  
 111…記録装置  
 112…クライアント  
 113…WWWページブラウザ  
 114…電子メールブラウザ  
 115…電子ニュースブラウザ  
 120…構造化文書処理部  
 121, 221…構造化文書要求解析部  
 122, 222…構造化文書受信部  
 123, 223…構造化文書内重要項目抽出部  
 124, 224…構造化文書内重要項目関連付加項目生成部  
 125, 225…構造化文書内重要項目指示制御部  
 126, 226…構造化文書内重要項目関連付加項目指示制御部  
 127, 227…検索指示制御部  
 128, 228…記憶部制御部  
 129, 229…記憶部  
 130, 230…構造化文書閲覧部  
 131, 231…構造化文書要求入力部  
 132, 232…構造化文書表示部  
 140, 240…重要項目指示部  
 141, 241…重要項目指示入力部  
 142, 242…重要項目指示実行結果表示部  
 150, 250…付加項目指示部  
 151, 251…付加項目指示入力部

152, 252…付加項目指示実行結果表示部  
 160, 260…検索指示部  
 161, 261…検索指示入力部  
 162, 262…検索結果表示部  
 170, 270…通信路  
 210…構造化文書閲覧装置  
 280…構造化文書送受信装置  
 501…類語データベース  
 502…抽出方法データベース  
 503…重要項目データベース  
 506…重要項目抽出方法設定アプリケーション  
 10201, 10301, 10601, 13001…ユーザ(ブラウザなど)  
 10401, 10501, 10701…ユーザインタフェース部  
 10202~10702, 13002…ユーザ認証部  
 10203~10703, 13003…構造化文書要求受信部  
 10204~10704, 13004…構造化文書送信部  
 10205~10705, 13005…重要項目抽出部  
 13051…共通重要項目処理部  
 10206~10706, 13006…構造化文書送受信部  
 10207~10707, 13007…重要項目格納部  
 10308, 10508~10708, 13008…アクセス許可判定部

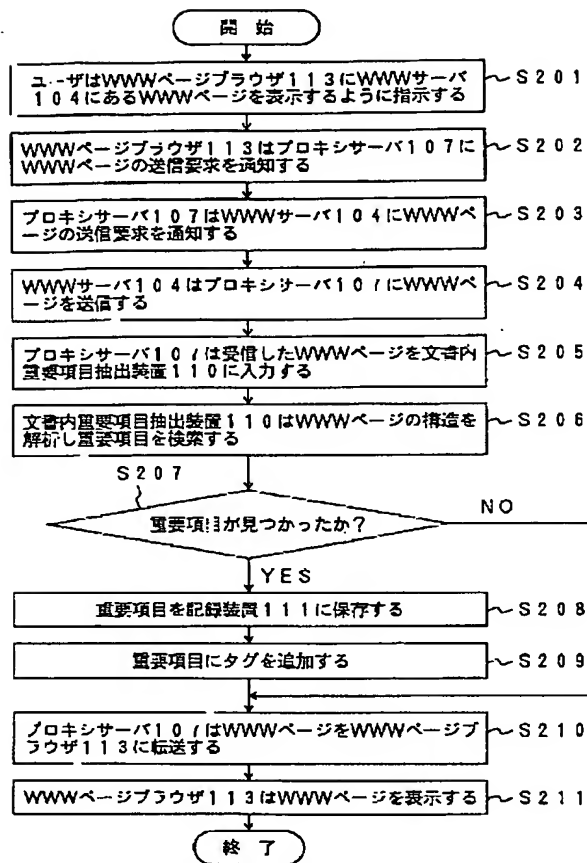
【図1】



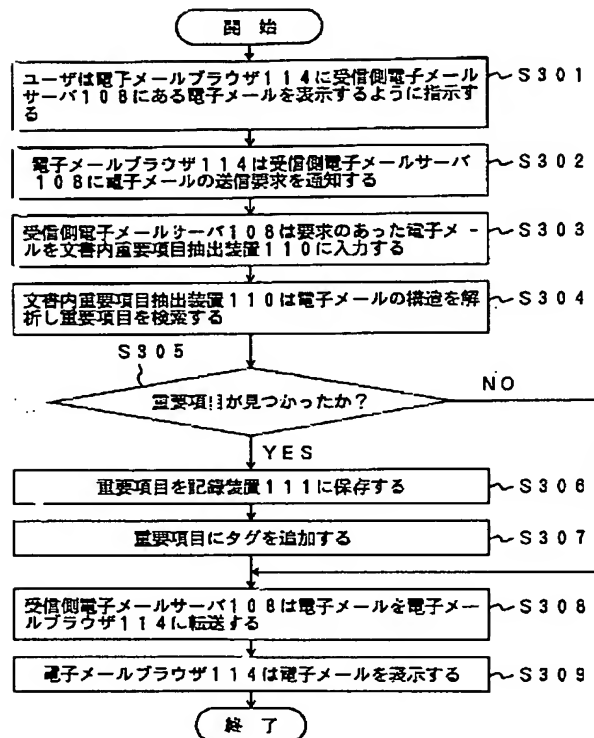
【図11】

類語	価格	値段	定価	特価	プライス
1101					

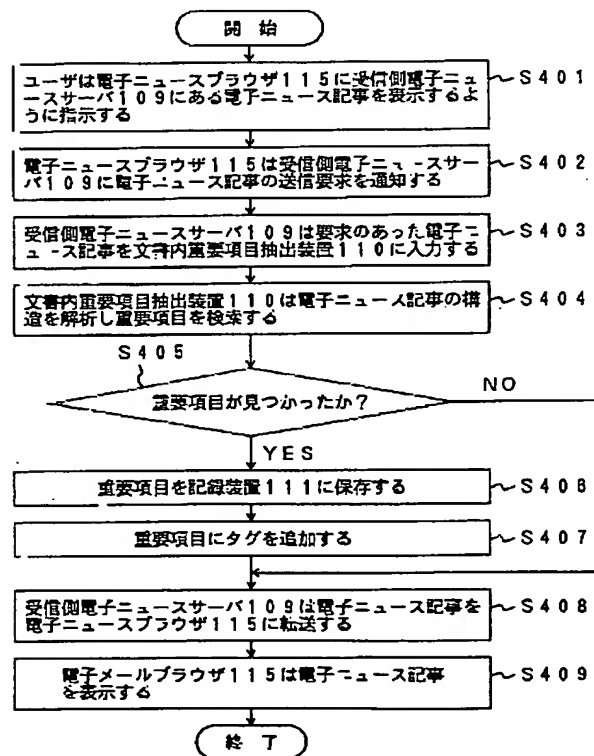
【図2】



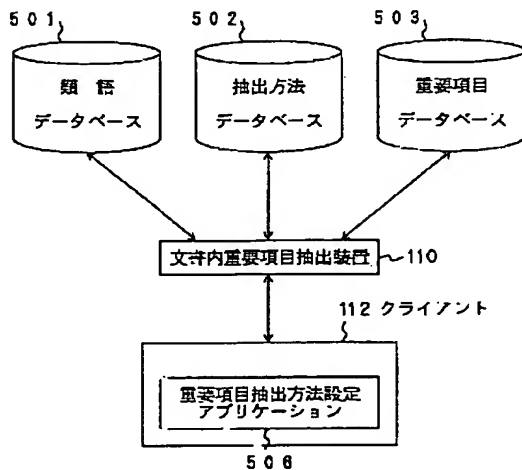
【図3】



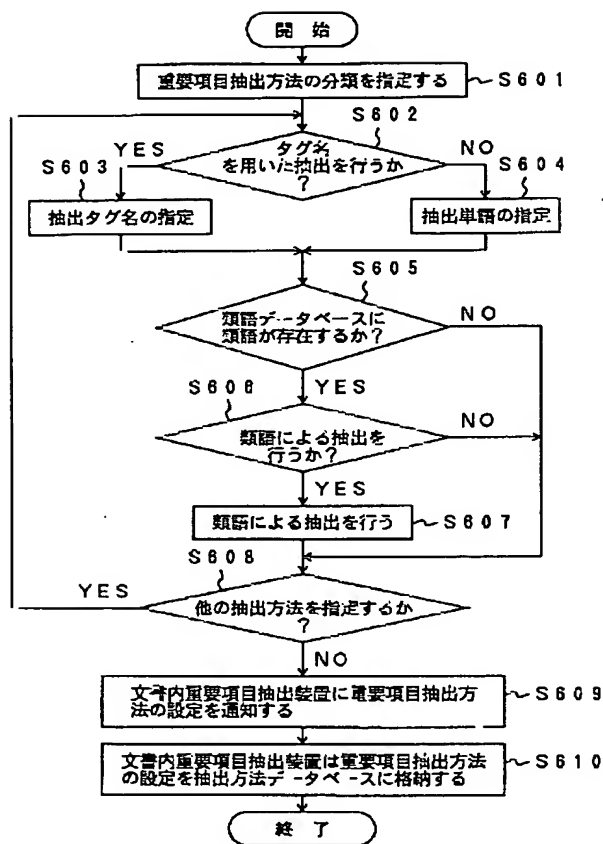
【図4】



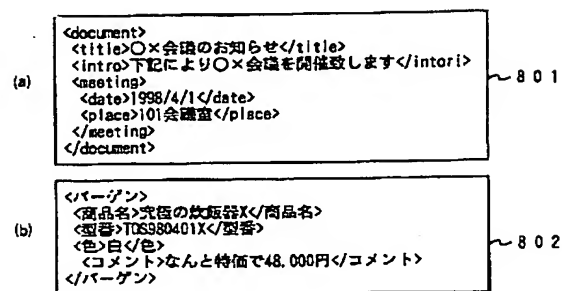
【図5】



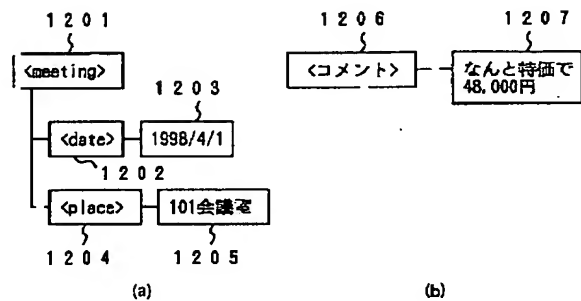
【図6】



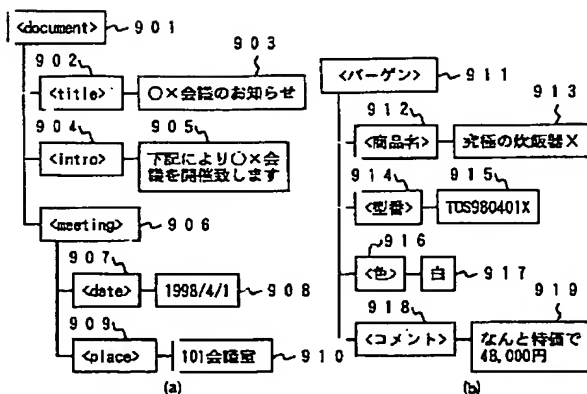
【図8】



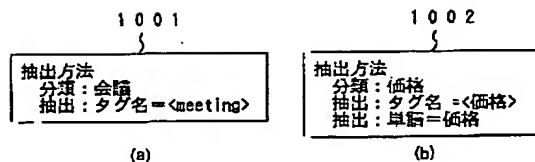
【図12】



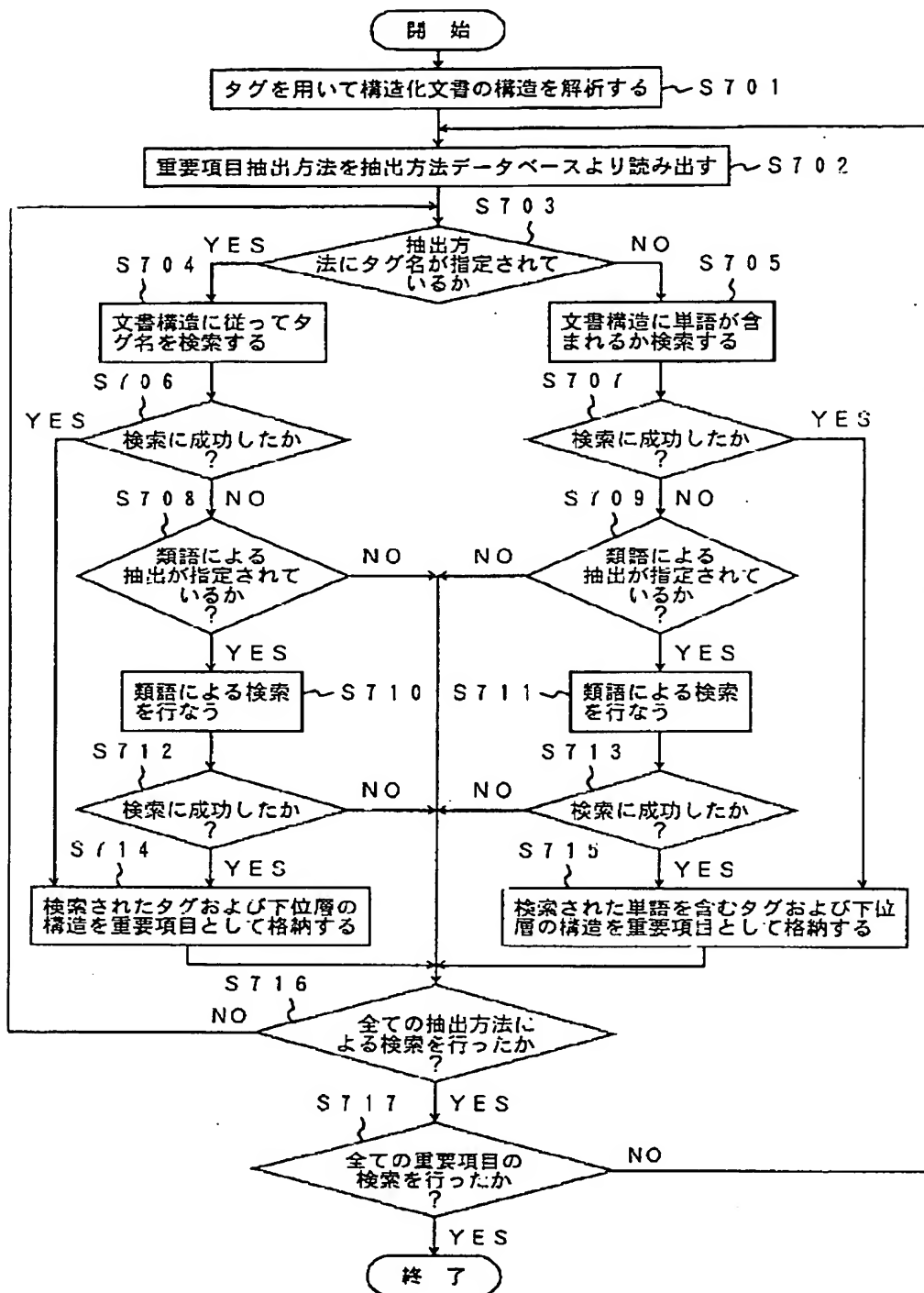
【図9】



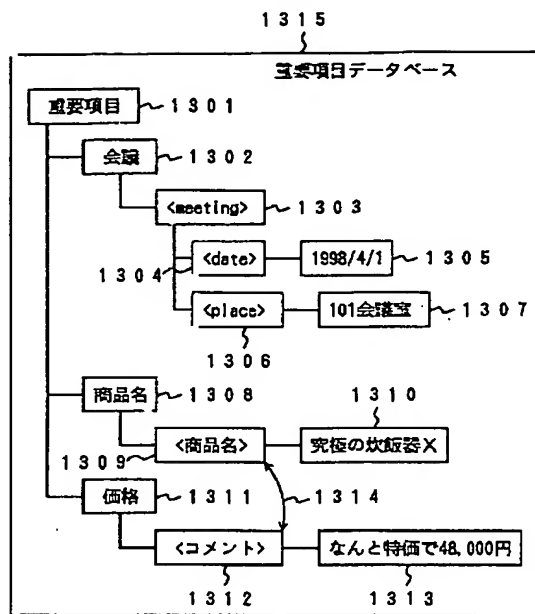
【図10】



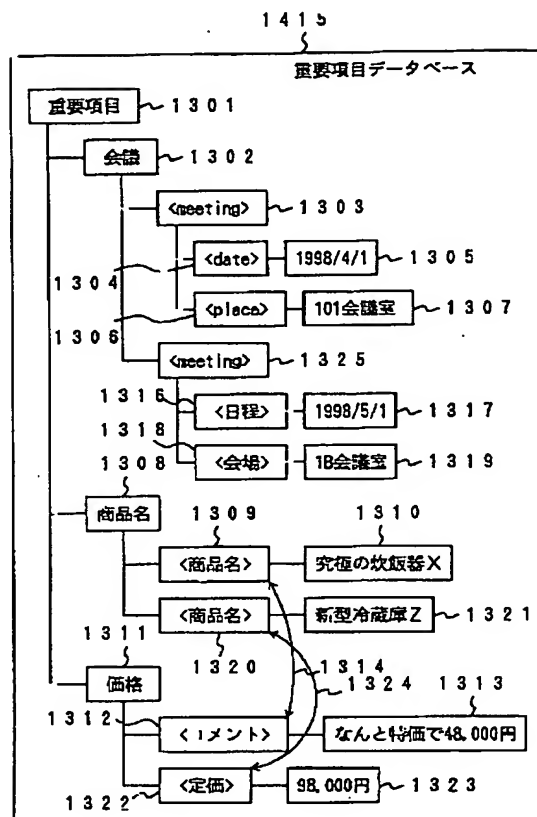
【 図 7 】



【図13】



【図14】



【図15】

文書表示

1601 会議のご案内 1602

拝啓...

敬具

記

日時 会場

指定 追加 変更 削除

1605 1606 1607 1608

【図16】

文書表示

1601 会議のご案内 1602

拝啓...

敬具

記

日時 会場

指定 追加 変更 削除

1605 1606 1607 1608



【図17】

重要項目の指定	
重要項目名	1801
1802	
抽出名	
1803 1804	
抽出方法	タグ 単語
1805 1806	
保存方法	全文 タグ
決定 1807	

【図18】

重要項目の追加	
重要項目名	2001
2002	
抽出名	
2003 2004	
抽出方法	タグ 単語
2005 2006	
保存方法	全文 タグ
決定 2007	

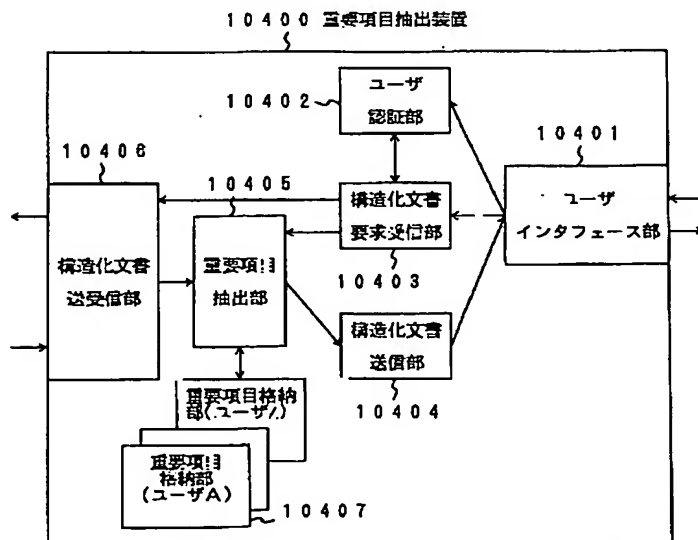
【図19】

重要項目の変更	
重要項目名	2201
2202	
抽出名	
2203 2204	
抽出方法	タグ 単語
2205 2206	
保存方法	全文 タグ
決定 2207	

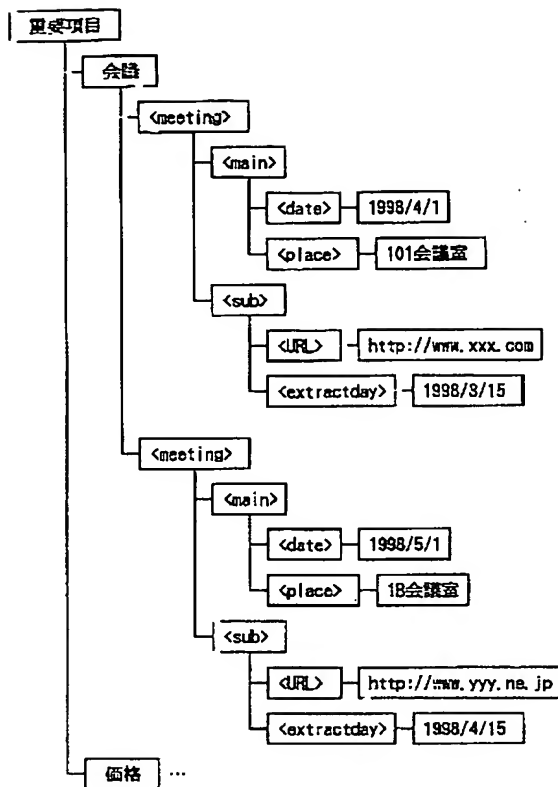
【図20】

重要項目の削除	
重要項目名	2401
2402	
抽出名	
2403 2404	
抽出方法	タグ 単語
2405 2406	
保存方法	全文 タグ
決定 2407	

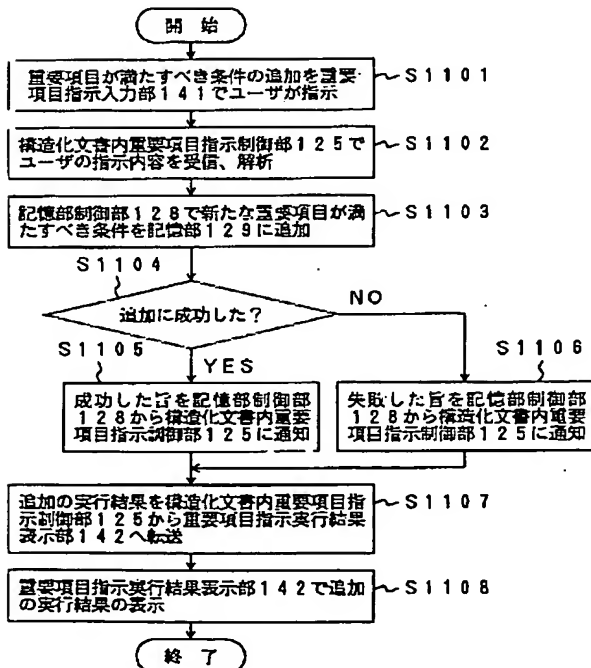
【図32】



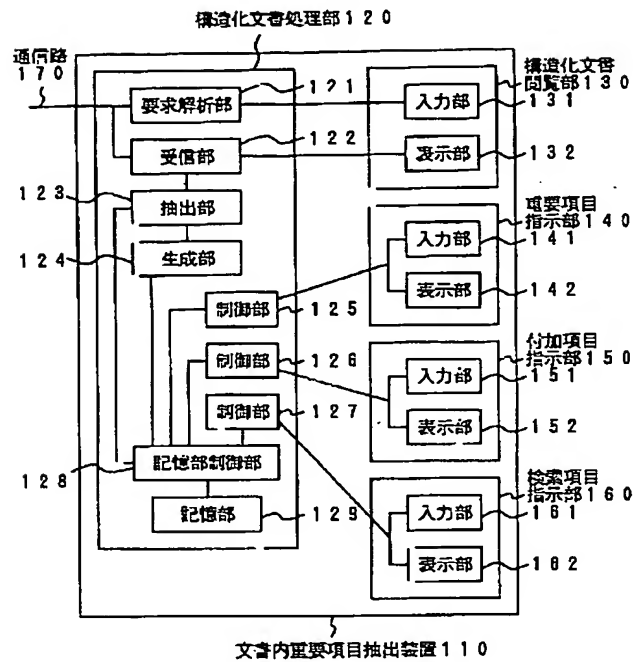
【図21】



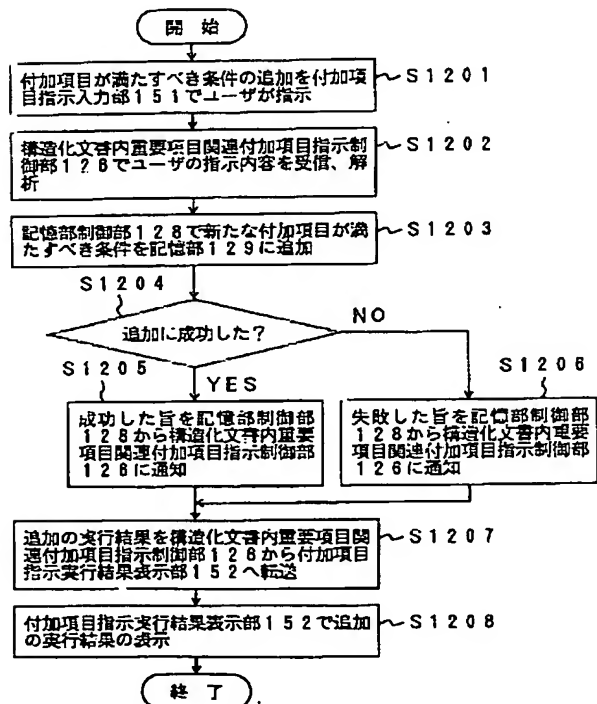
【図23】



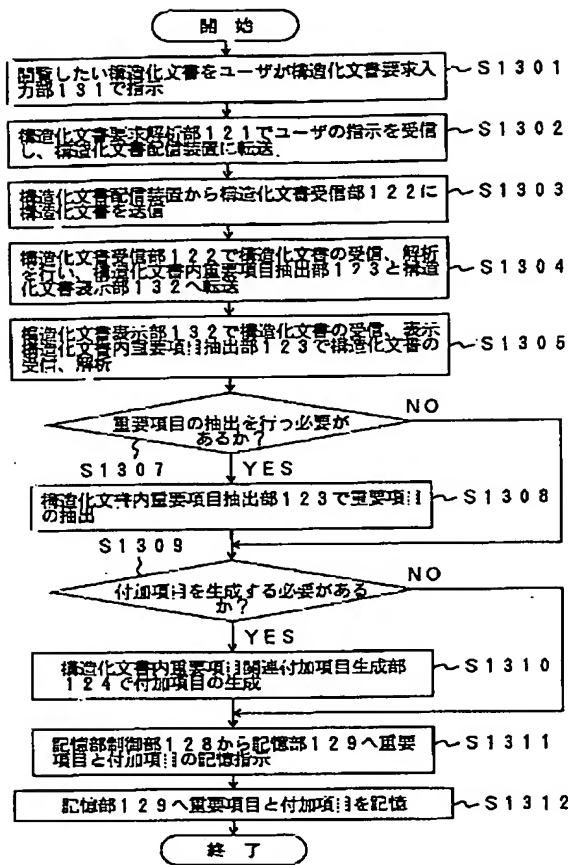
【図22】



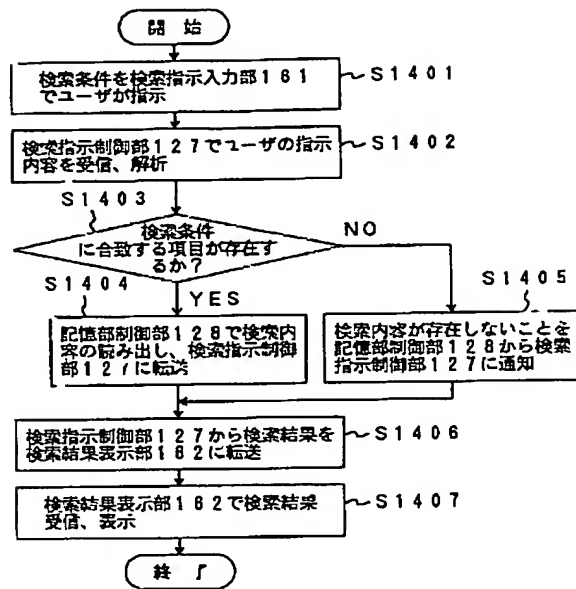
【図24】



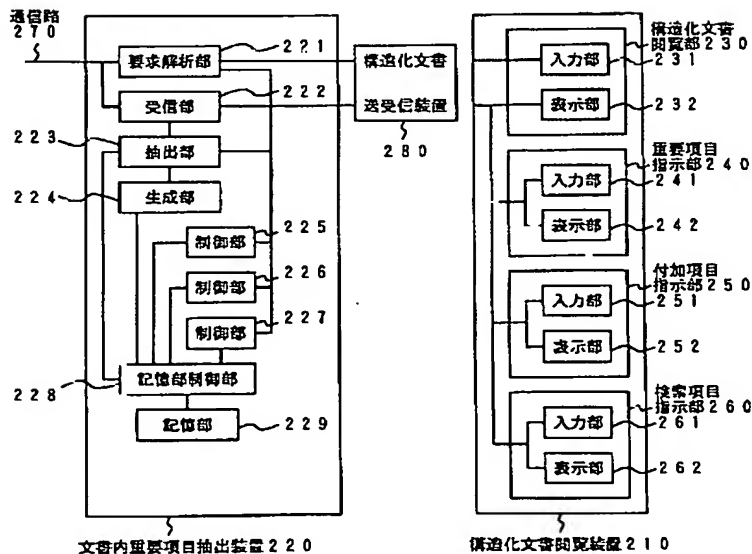
【図25】



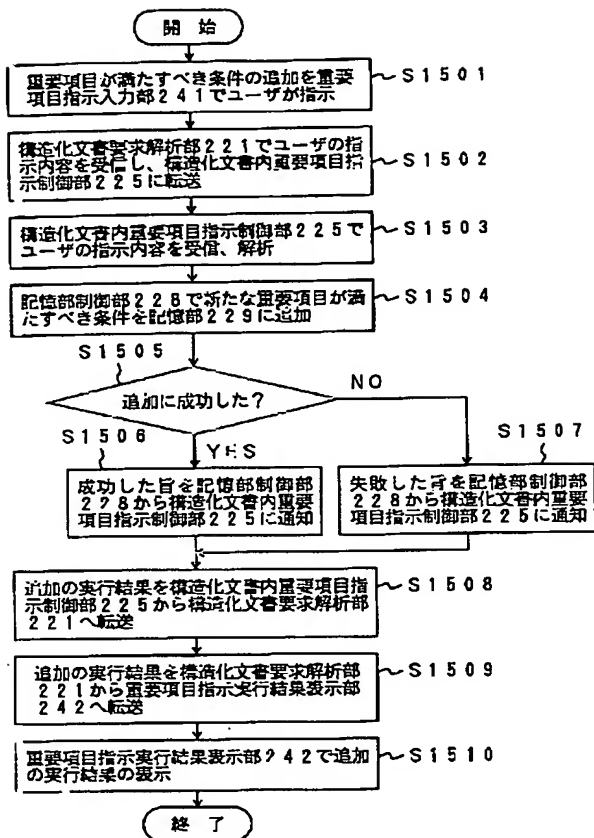
【図26】



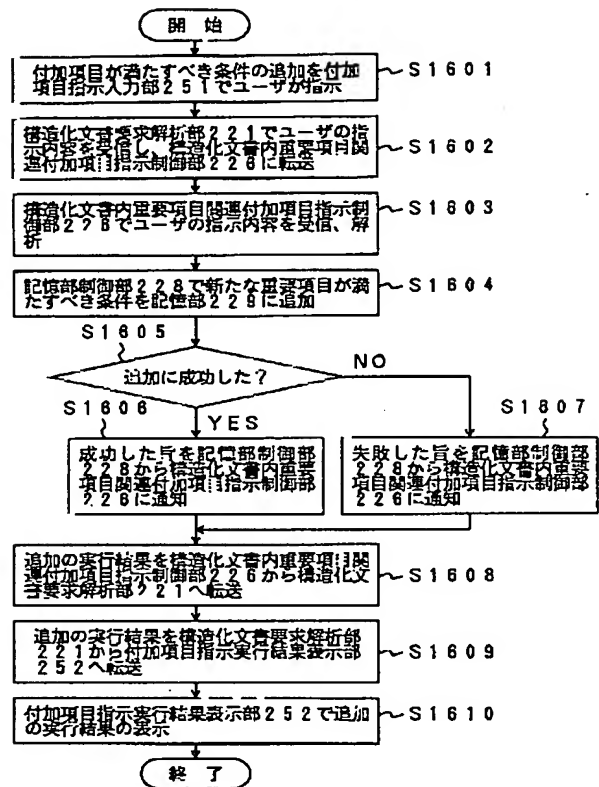
【図27】



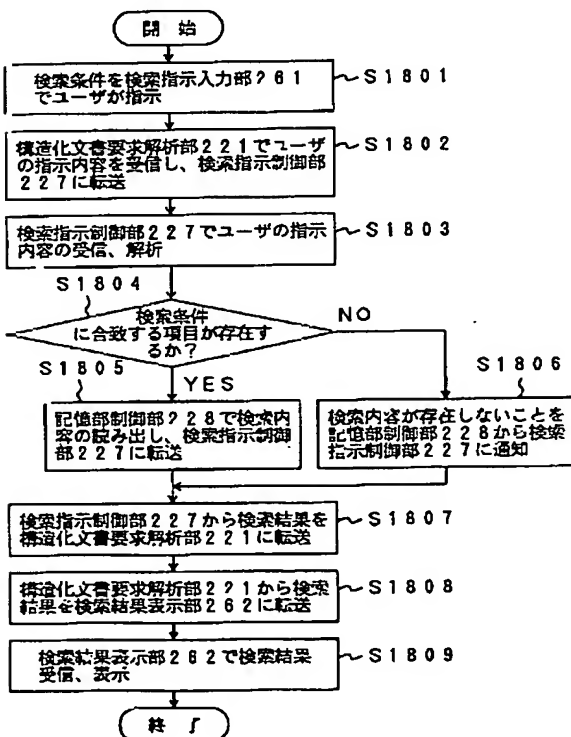
【図28】



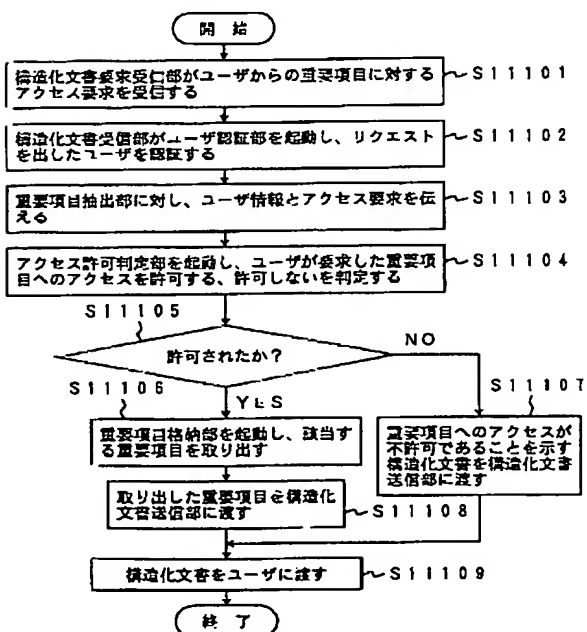
【図29】



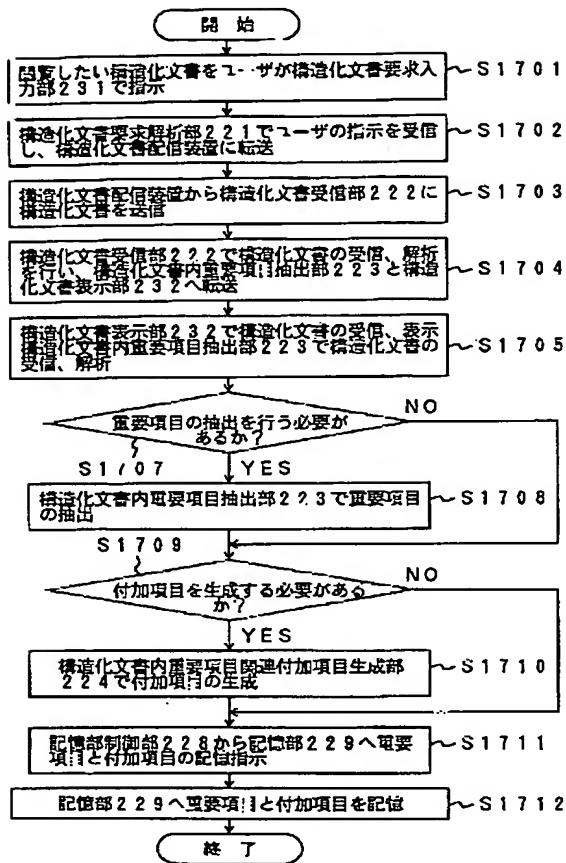
【図31】



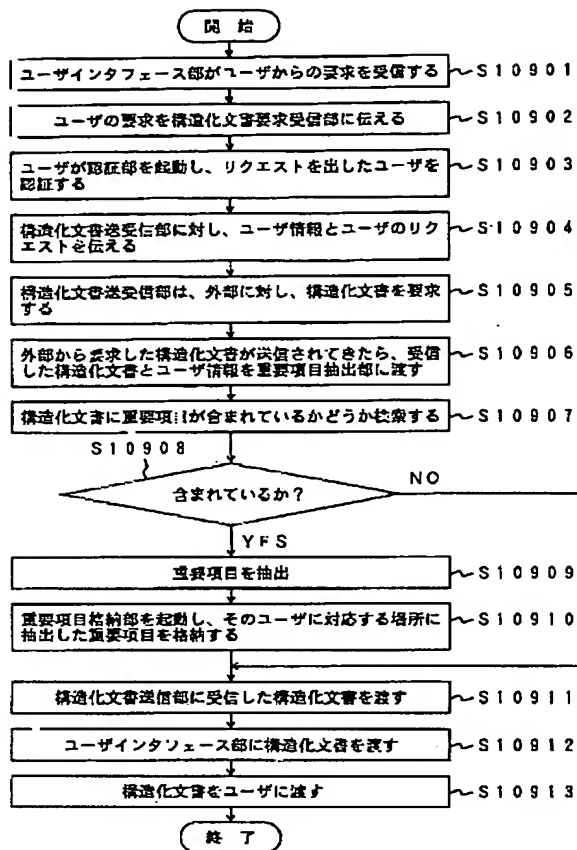
【図42】



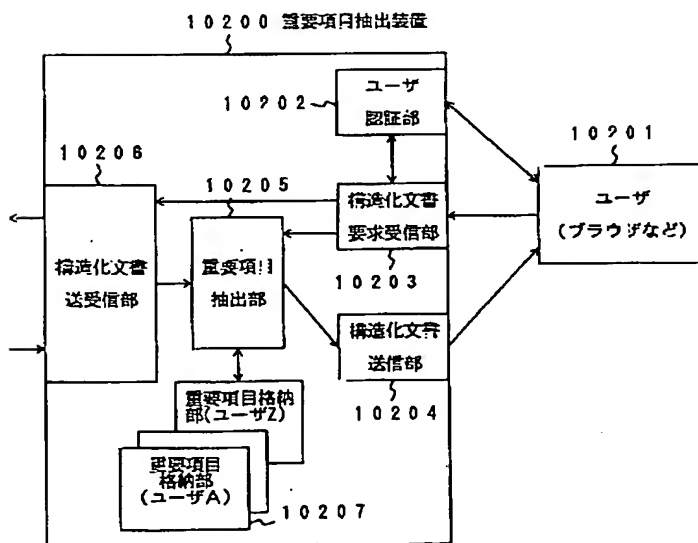
【図30】



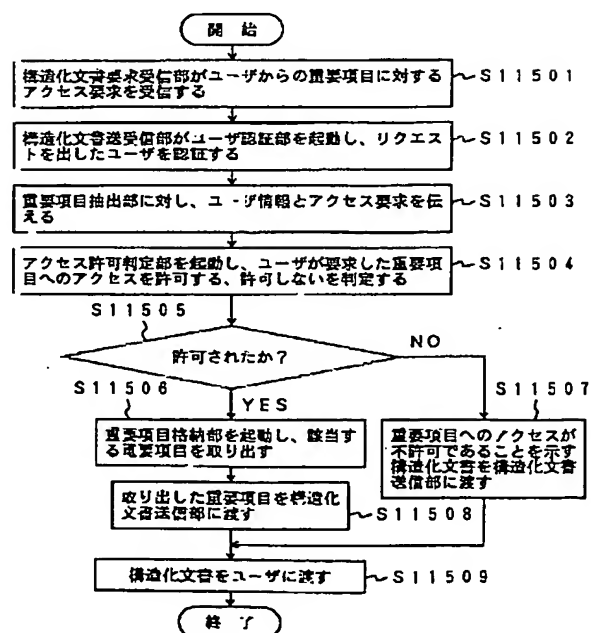
【図33】



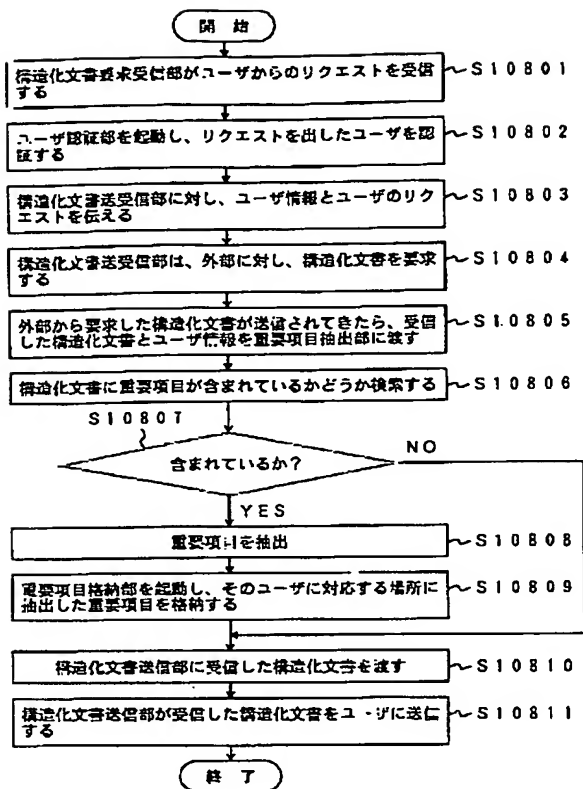
【図34】



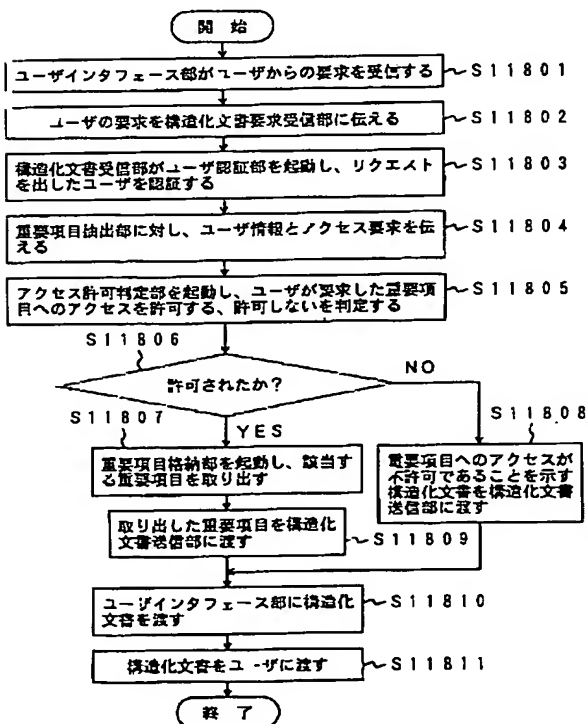
【図49】



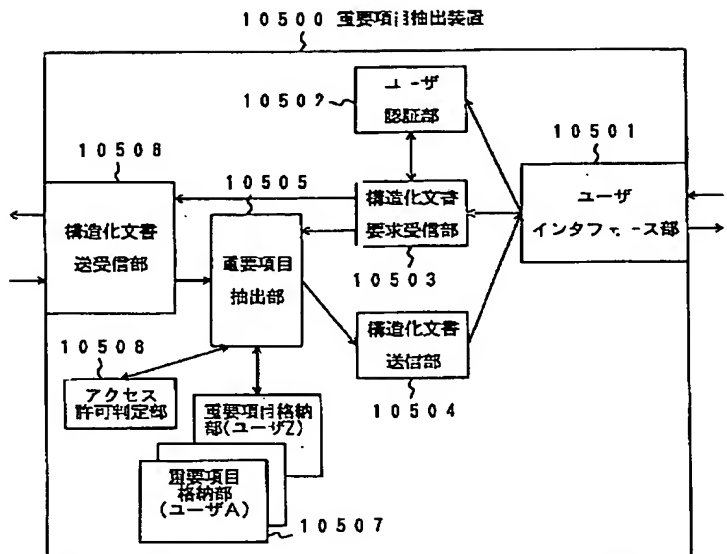
【図35】



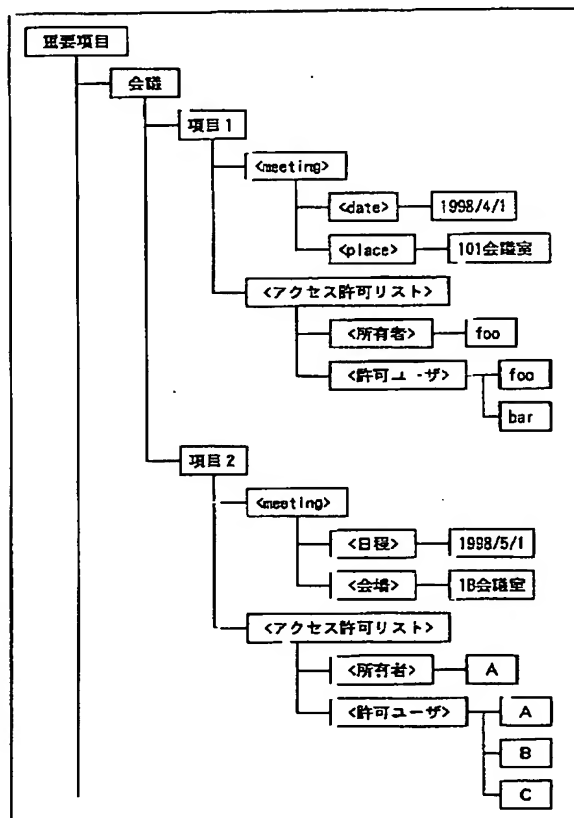
【図45】



【図36】

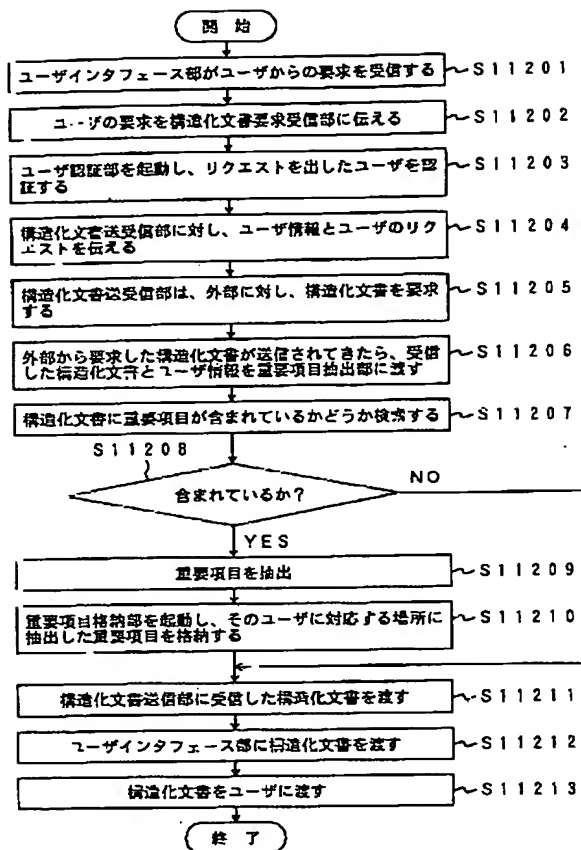


【図37】

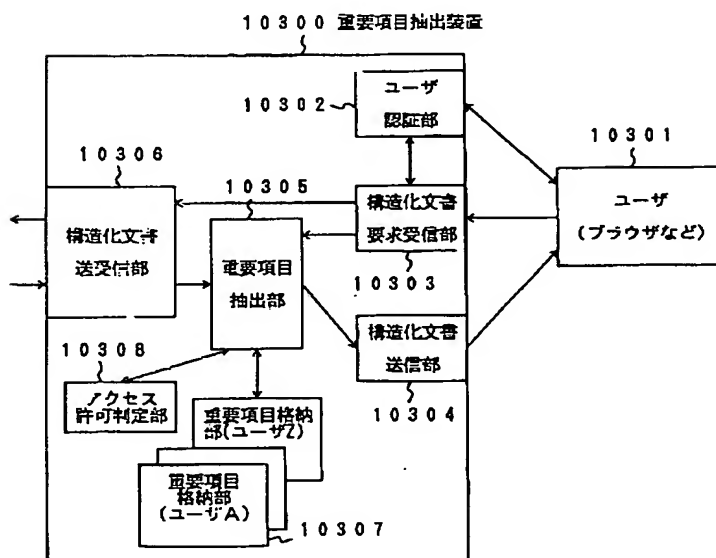




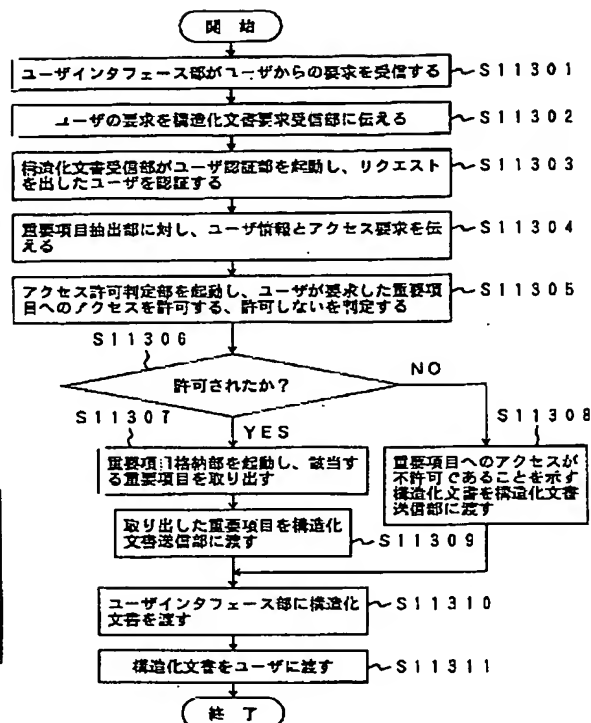
【図38】



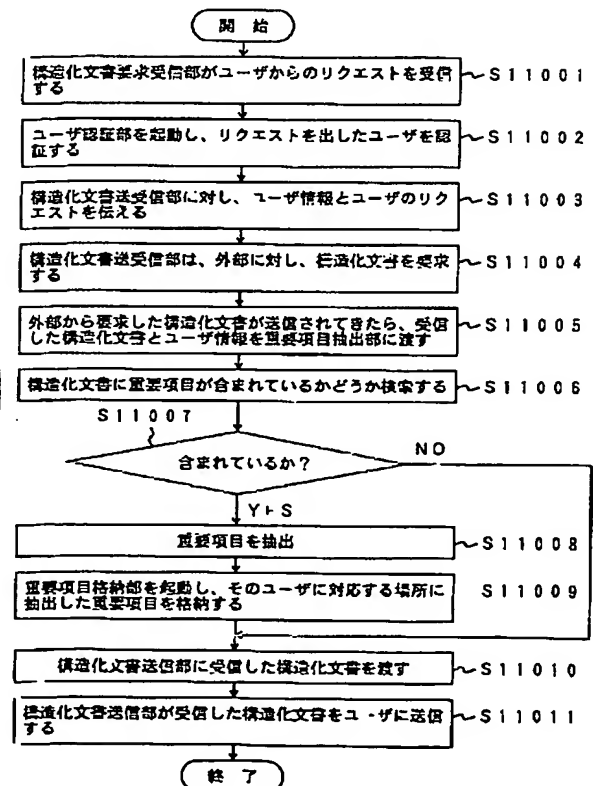
【図40】



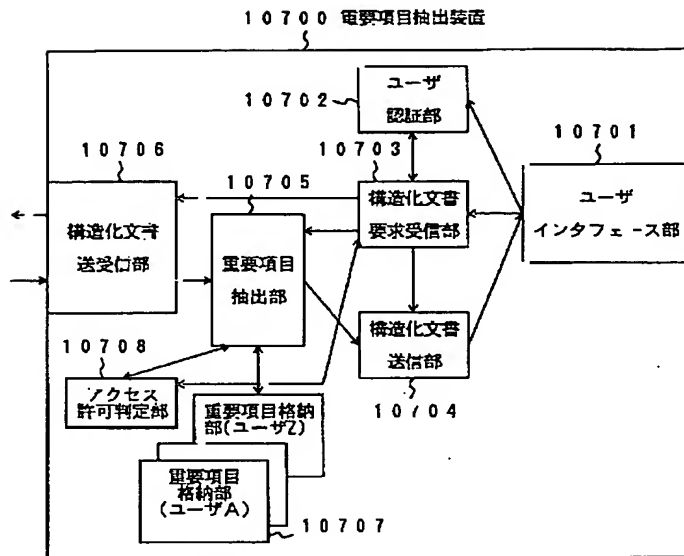
【図39】



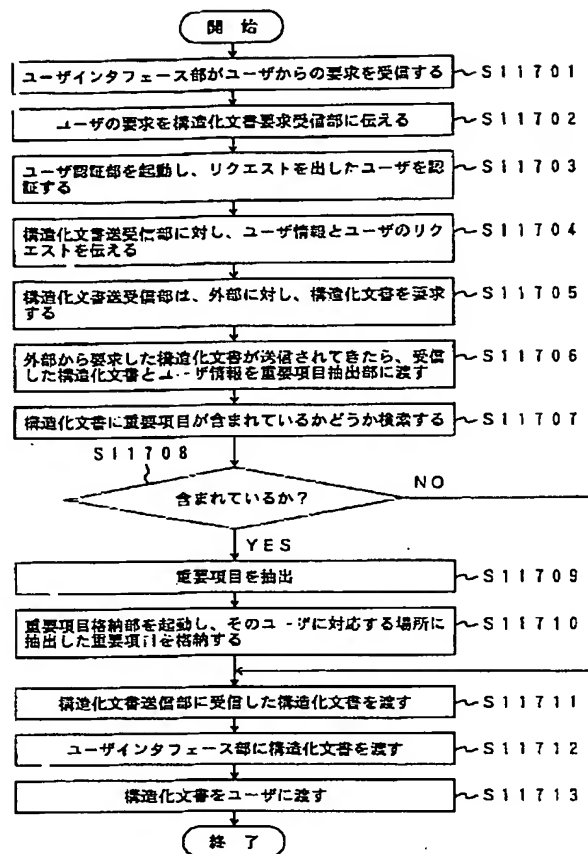
【図41】



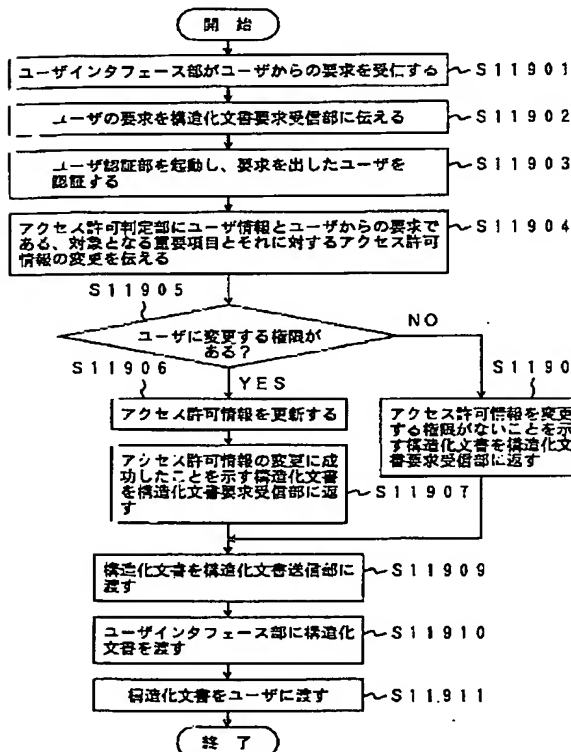
【図43】



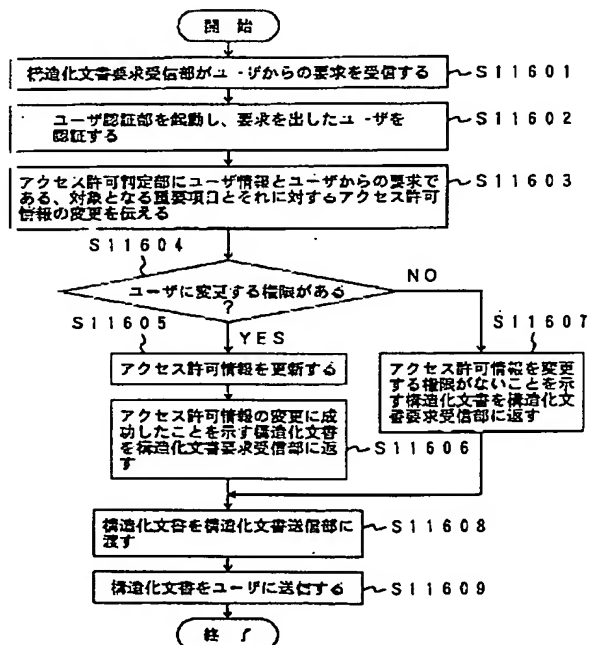
【図44】



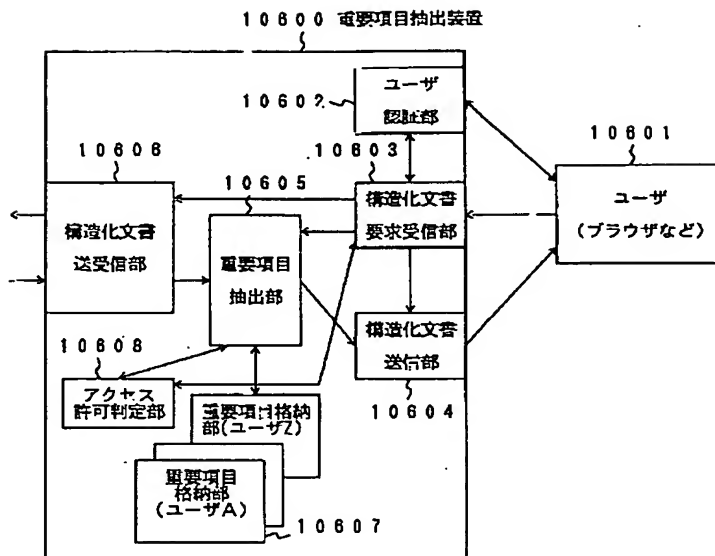
【図46】



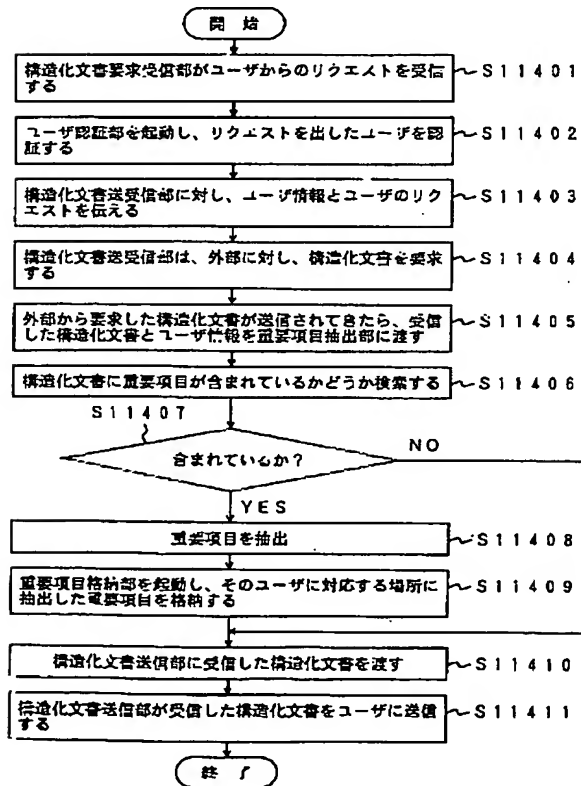
【図50】



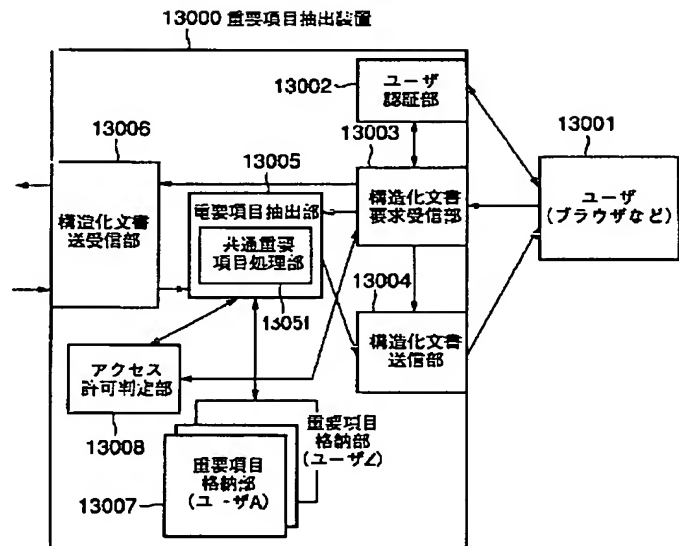
【図47】



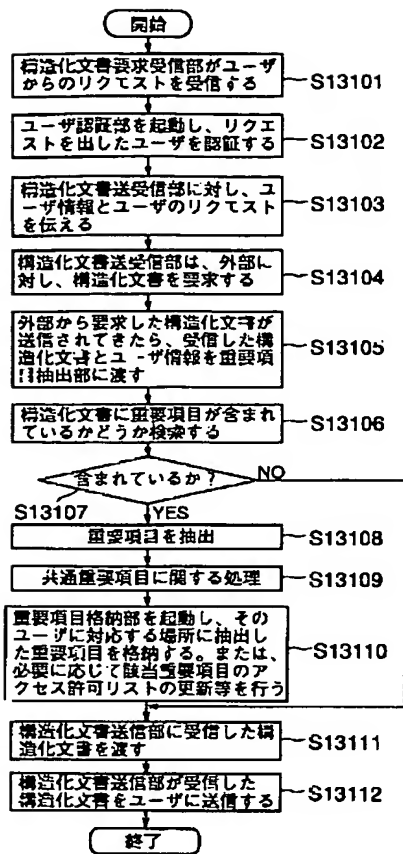
【図48】



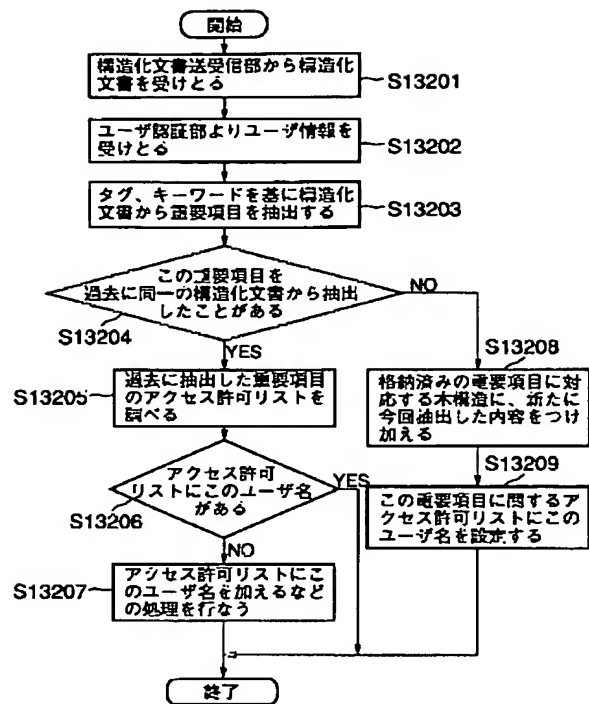
【図51】



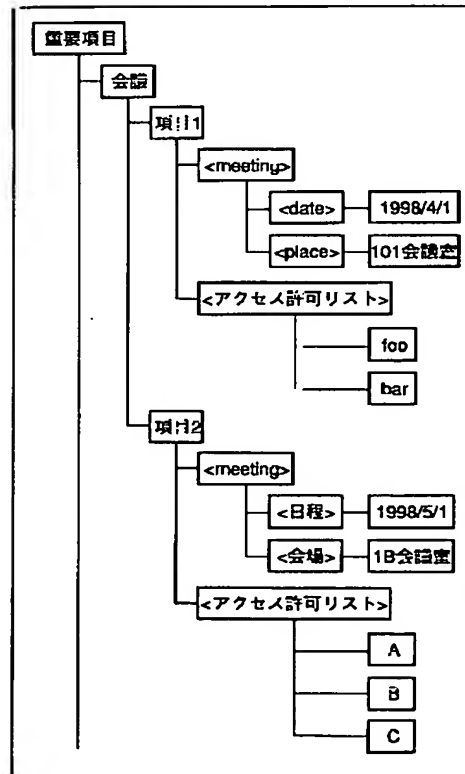
【図52】



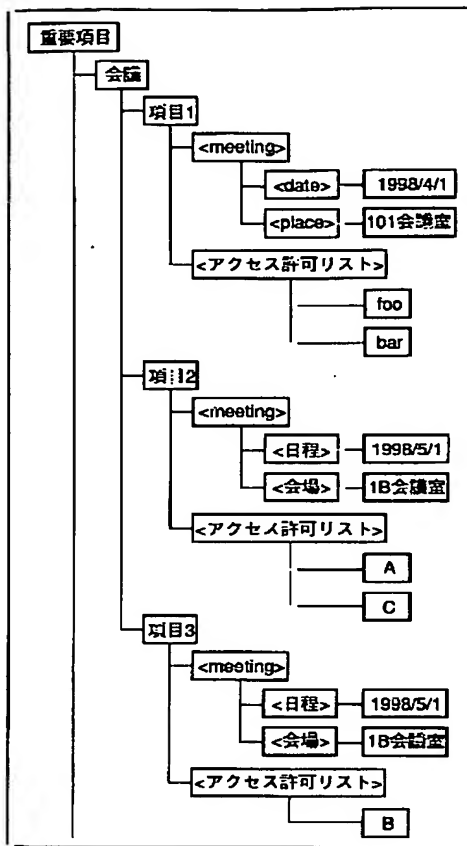
【図53】



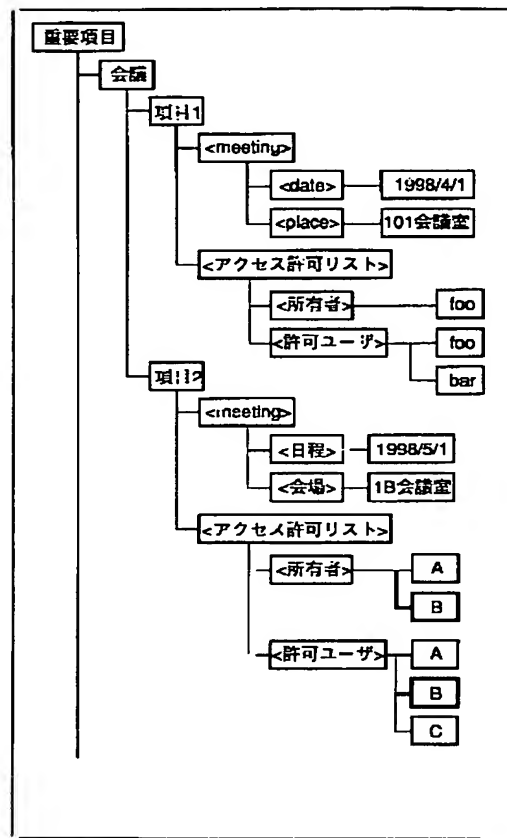
【図54】



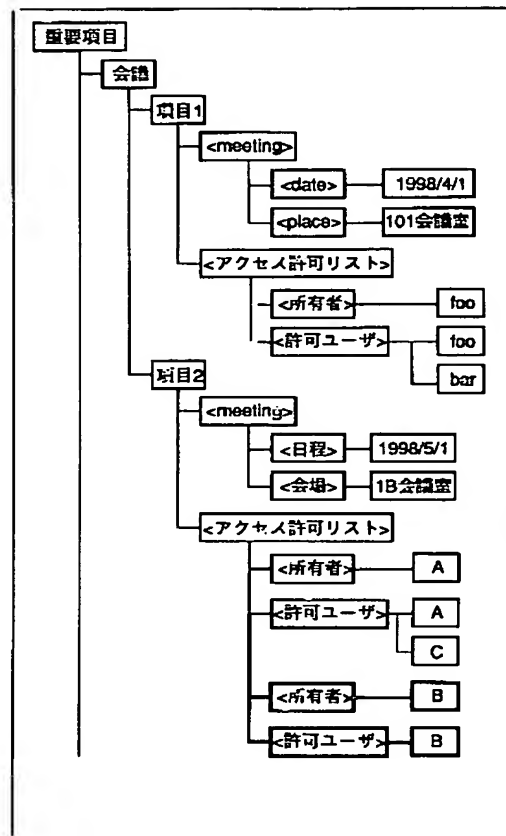
【図55】



【図56】



【図57】



フロントページの続き

(72)発明者 鳥井 修  
神奈川県川崎市幸区小向東芝町1番地 株  
式会社東芝研究開発センター内  
(72)発明者 金井 達徳  
神奈川県川崎市幸区小向東芝町1番地 株  
式会社東芝研究開発センター内  
(72)発明者 横川 健  
神奈川県川崎市幸区小向東芝町1番地 株  
式会社東芝研究開発センター内

(72)発明者 矢尾 浩  
神奈川県川崎市幸区小向東芝町1番地 株  
式会社東芝研究開発センター内  
(72)発明者 田中 久子  
神奈川県川崎市幸区小向東芝町1番地 株  
式会社東芝研究開発センター内  
Fターム(参考) 5B009 QA03  
5B075 KK34 KK39 ND03 ND34 NK02  
PQ02 PQ32 UU06